Classical capacity of quantum channels with general Markovian correlated noise

Nilanjana Datta*& Tony Dorlas†

Statistical Laboratory

Centre for Mathematical Sciences

University of Cambridge

Wilberforce Road, Cambridge CB3 0WB, U.K.

December 5, 2007

Abstract

The classical capacity of a quantum channel with *arbitrary* Markovian correlated noise is evaluated. For an irreducible and aperiodic Markov Chain, the channel is forgetful, and one retrieves the known expression [15] for the capacity. For the more general case of a channel with long-term memory, which corresponds to a Markov chain which does not converge to equilibrium, the capacity is expressed in terms of the communicating classes of the Markov chain.

*email: n.datta@statslab.cam.ac.uk.

†email:dorlas@stp.dias.ie; on leave from the Dublin Institute of Advanced Studies, School of Theoretical Physics, 10 Burlington Road, Dublin 4, Ireland

1 Introduction

Shannon, in his celebrated Noisy Channel Coding Theorem [22], obtained an explicit expression for the channel capacity of discrete, memoryless[1], classical channels. The first rigorous proof of this fundamental theorem was provided by Feinstein [8]. He used a packing argument (see e.g.[10]) to find a lower bound to the maximal number of codewords that can be sent through the channel reliably, i.e., with an arbitrarily low probability of error. More precisely, he proved that for any given $\delta > 0$, and sufficiently large number, n, of uses of a memoryless classical channel, the lower bound to the maximal number, N_n, of codewords that can be transmitted through the channel reliably, is given by

$$N_n \geq 2^{n(H(X:Y)-\delta)}.$$

Here $H(X : Y)$ is the mutual information of the random variables X and Y, corresponding to the input and the output of the channel, respectively. This lower bound implies that for n large enough, any real number $R < C = \max H(X : Y)$, (the maximum being taken over all possible input distributions), at least $N_n = [2^{nR}]$ classical messages can be transmitted through the channel reliably. In other words, any rate $R < C$ is *achievable*.

The assumption that noise is uncorrelated between successive uses of a channel is not realistic. Hence memory effects need to be taken into account. In this paper we consider the transmission of classical information through a class of quantum channels with memory. The first model of such a channel was studied by Macchiavello and Palma [17]. They showed that the transmission of classical information through two successive uses of a quantum depolarising channel, with Markovian correlated noise, is enhanced by using inputs entangled over the two uses. A more general model of a quantum channel with memory was introduced by Bowen and Mancini [4] and also studied by Kretschmann and Werner [15]. In particular, in [15], the capacities of a class of quantum channels with memory, the so-called *forgetful channels* were evaluated. Similar results were obtained by Bjelaković and Boche [2].

[1]For such a channel, the noise affecting successive input states, is assumed to be perfectly uncorrelated.

Further, in [7], the classical capacity of a class of quantum channels with long-term memory was obtained. The memory of the channel considered in [7] can be viewed as a special case of a general Markovian memory, where the Markov chain is aperiodic but not irreducible, and hence does not converge to equilibrium. Recently, there was a generalization of the result of [7] by Bjelaković and Boche, who in [3] obtained the classical capacities of compound and averaged quantum channels.

Another interesting special case of a channel with long-term memory is that in which the memory is described by a periodic Markov chain. A simple example of this is a channel given by alternating applications of two completely positive trace preserving (CPT) maps Φ_1 and Φ_2, with the first map being Φ_1 or Φ_2 with probability $1/2$.

In this paper we study channels with *arbitrary* Markovian correlated noise. This includes, in particular, the above special cases. We show that the capacity in the general case can be expressed in terms of the communicating classes of the underlying Markov chain.

We start the main body of our paper with some preliminaries in Section 2. In Section 3, the quantum channel is defined and its capacity is stated in the main theorem, Theorem 1, of this paper. In Section 4, we prove a special case of the direct part of this theorem, corresponding to a Markov chain which converges to equilibrium and is hence forgetful. This section therefore provides an alternative proof of the result of Kretschmann and Werner [15] for the classical capacity of such a channel. This proof is extended to the case of an arbitrary Markov chain in Section 5. In the latter, we employ the idea of adding a preamble to the codewords (as was done in [7]) in order to distinguish between the different communicating classes of the Markov Chain. The proof of the (weak) converse part of our main result (Theorem 1) is given in Section 6.

2 Mathematical Preliminaries

Let $\mathcal{H}$ and $\mathcal{K}$ be given finite-dimensional Hilbert spaces and denote by $\mathcal{B}(\mathcal{H})$ the algebra of linear operators on $\mathcal{H}$. We also consider the tensor product algebras $\mathcal{A}_n = \mathcal{B}(\mathcal{H}^{\otimes n})$ and the infinite tensor product C*-algebra obtained as the strong closure

$$\mathcal{A}_\infty = \overline{\bigcup_{n=1}^{\infty} \mathcal{A}_n}, \tag{1}$$

where we embed $\mathcal{A}_n$ into $\mathcal{A}_{n+1}$ in the obvious way. Similarly, we define $\mathcal{B}_n = \mathcal{B}(\mathcal{K}^{\otimes n})$ and $\mathcal{B}_\infty$. A *state* on an algebra $\mathcal{A}$ is a positive linear functional ϕ on $\mathcal{A}$ with $\phi(\mathbf{1}) = 1$, where $\mathbf{1}$ denotes identity operator. If $\mathcal{A}$ is finite-dimensional then there exists a density matrix ρ_ϕ (i.e., a positive operator with $\mathrm{Tr}\,\rho_\phi = 1$) such that $\phi(A) = \mathrm{Tr}\,(\rho_\phi A)$, for any $A \in \mathcal{A}$. We denote the states on $\mathcal{A}_\infty$ by $\mathcal{S}(\mathcal{A}_\infty)$, those on $\mathcal{A}_n$ by $\mathcal{S}(\mathcal{A}_n)$,etc.

3 A quantum channel with classical memory

Let there be given a Markov chain on a finite state space I with transition probabilities $\{q_{ii'}\}_{i,i'\in I}$ and let $\{\gamma_i\}_{i\in I}$ be an invariant distribution for this chain, i.e.

$$\gamma_{i'} = \sum_{i\in I} \gamma_i q_{ii'}. \tag{2}$$

Moreover, let $\Phi_i : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ be given completely positive trace-preserving (CPT) maps for each $i \in I$. Then we define a quantum channel with Markovian correlated noise, by the CPT map $\Phi_\infty : \mathcal{S}(\mathcal{A}_\infty) \to \mathcal{S}(\mathcal{B}_\infty)$ on the states of $\mathcal{A}_\infty$ by

$$(\Phi_\infty)(\phi)(A) = \sum_{i_1,\dots,i_n \in I} \gamma_{i_1} q_{i_1 i_2} \dots q_{i_{n-1} i_n} \mathrm{Tr}\left[(\Phi_{i_1} \otimes \cdots \otimes \Phi_{i_n})(\rho_{\phi_n})\, A\right] \tag{3}$$

for $A \in \mathcal{B}_n$. Here, ϕ_n is the restriction of ϕ to $\mathcal{A}_n$ and ρ_{ϕ_n} its density matrix. It is easily seen, using the property (2), that this definition is consistent and defines a CPT map on the states of $\mathcal{A}_\infty$, and moreover, that it is translation-invariant (stationary).

We denote the transpose action of the restriction of Φ_∞ to $\mathcal{S}(\mathcal{A}_n)$ by $\Phi^{(n)} : \mathcal{B}(\mathcal{H}^{\otimes n}) \to \mathcal{B}(\mathcal{K}^{\otimes n})$, i.e.,

$$\mathrm{Tr}\left(\Phi^{(n)}(\rho_\phi)A\right) = (\Phi_\infty(\phi))(A),$$

for a density matrix $\rho_\phi \in \mathcal{B}(\mathcal{H}^{\otimes n})$, $\phi \in \mathcal{S}(\mathcal{A}_n)$.

Note that

$$\Phi^{(n)}(\rho^{(n)}) = \sum_{i_1,\ldots,i_n \in I} \gamma_{i_1} q_{i_1 i_2} \ldots q_{i_{n-1} i_n} (\Phi_{i_1} \otimes \cdots \otimes \Phi_{i_n})(\rho^{(n)}). \tag{4}$$

Let us consider the transmission of classical information through $\Phi^{(n)}$. Suppose Alice has a set of messages, labelled by the elements of the set $\mathcal{M}_n = \{1, 2, \ldots, M_n\}$, which she would like to communicate to Bob, using the quantum channel Φ. To do this, she encodes each message into a quantum state of a physical system with Hilbert space $\mathcal{H}^{\otimes n}$, which she then sends to Bob through n uses of the quantum channel. In order to infer the message that Alice communicated to him, Bob makes a measurement (described by POVM elements) on the state that he receives. The encoding and decoding operations, employed to achieve reliable transmission of information through the channel, together define a quantum error correcting code (QECC). More precisely, a code $\mathcal{C}^{(n)}$ of size N_n is given by a sequence $\{\rho_i^{(n)}, E_i^{(n)}\}_{i=1}^{N_n}$ where each $\rho_i^{(n)}$ is a state in $\mathcal{B}(\mathcal{H}^{\otimes n})$ and each $E_i^{(n)}$ is a positive operator acting in $\mathcal{K}^{\otimes n}$, such that $\sum_{i=1}^{N_n} E_i^{(n)} \leq \mathbf{1}^{(n)}$. Here, $\mathbf{1}^{(n)}$ denotes the identity operator in $\mathcal{B}(\mathcal{K}^{\otimes n})$. Defining $E_0^{(n)} = \mathbf{1}^{(n)} - \sum_{i=1}^{N_n} E_i^{(n)}$, yields a Positive Operator-Valued Measure (POVM) $\{E_i^{(n)}\}_{i=0}^{N_n}$ in $\mathcal{K}^{\otimes n}$. An output $i \geq 1$ would lead to the inference that the state (or codeword) $\rho_i^{(n)}$ was transmitted through the channel $\Phi^{(n)}$, whereas the output 0 is interpreted as a failure of any inference. The average probability of error for the code $\mathcal{C}^{(n)}$ is given by

$$P_e(\mathcal{C}^{(n)}) := \frac{1}{N_n} \sum_{i=1}^{N_n} \left(1 - \mathrm{Trace}\big(\Phi^{(n)}(\rho_i^{(n)}) E_i^{(n)}\big)\right), \tag{5}$$

If there exists an $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, there exists a sequence of codes $\{\mathcal{C}^{(n)}\}_{n=1}^{\infty}$, of sizes $N_n \geq 2^{nR}$, for which $P_e(\mathcal{C}^{(n)}) \to 0$ as $n \to \infty$, then R is said to be an *achievable* rate.

The *classical capacity* of Φ is defined as

$$C(\Phi) := \sup R, \tag{6}$$

where R is an achievable rate.

Let $\mathcal{C}$ be the set of communicating classes, C, of the Markov chain [19] for which

$$\gamma_C = \sum_{i \in C} \gamma_i > 0. \tag{7}$$

Any other classes can be disregarded. For $C \in \mathcal{C}$ we define

$$\Phi_C^{(n)}(\rho^{(n)}) := \frac{1}{\gamma_C} \sum_{i_1,\dots,i_n \in C} \gamma_{i_1} q_{i_1 i_2} \dots q_{i_{n-1} i_n} (\Phi_{i_1} \otimes \dots \otimes \Phi_{i_n})(\rho^{(n)}), \tag{8}$$

which represents the restriction of the classical memory of the channel to the class C. Notice that the Markov chain restricted to $C \in \mathcal{C}$ is necessarily irreducible, and is either aperiodic or periodic with a single period. In fact,

$$\mathcal{C} = \mathcal{C}_{aper} \cup \mathcal{C}_{per},$$

where $\mathcal{C}_{aper}$ denotes the set of communicating classes in $\mathcal{C}$ which are aperiodic, while $\mathcal{C}_{per}$ denotes the set of communicating classes in $\mathcal{C}$ which are aperiodic.

If $C \in \mathcal{C}_{aper}$, we define, for any ensemble $\{p_j^{(n)}, \rho_j^{(n)}\}$ of states on $\mathcal{H}^{\otimes n}$, the *mean Holevo quantity* for the class C as

$$\bar{\chi}_C^{(n)}(\{p_j^{(n)}, \rho_j^{(n)}\}) := \frac{1}{n} \left[S\left(\sum_j p_j^{(n)} \Phi_C^{(n)}(\rho_j^{(n)}) \right) - \sum_j p_j^{(n)} S\left(\Phi_C^{(n)}(\rho_j^{(n)}) \right) \right]. \tag{9}$$

If $C \in \mathcal{C}_{per}$ is periodic, with period L, then $C = \{i_0, i_1, \dots, i_{L-1}\}$ for certain $i_0, \dots, i_{L-1} \in I$, and $q_{i_k i_{k+1}} = 1$ for $k = 0, \dots, L-2$ and $q_{i_{L-1} i_0} = 1$. In this case,

$$\gamma_i = \frac{1}{L} \gamma_C \quad (i \in C) \tag{10}$$

and we set

$$\bar{\chi}_C^{(n)}(\{p_j^{(n)}, \rho_j^{(n)}\}) = \frac{1}{nL} \sum_{i \in C} \chi_{C,i}^{(n)}(\{p_j^{(n)}, \rho_j^{(n)}\}), \tag{11}$$

where for $k \in \{0, 1, \dots, L-1\}$,

$$\chi_{C,i_k}^{(n)}(\{p_j^{(n)}, \rho_j\}) = S\left(\Phi_{i_k} \otimes \Phi_{i_{k+1}} \dots \otimes \Phi_{i_{k+n-1}}(\bar{\rho}^{(n)})\right) - \bar{S}_{i_k}^{(n)}, \tag{12}$$

(the indices in the subscripts being taken modulo L), with

$$\bar{\rho}^{(n)} = \sum_j p_j^{(n)} \rho_j^{(n)}, \text{ and } \bar{S}_{i_k}^{(n)} = \sum_j p_j^{(n)} S(\Phi_{i_k} \otimes \Phi_{i_{k+1}} \cdots \otimes \Phi_{i_{k+n-1}}(\rho_j^{(n)})). \quad (13)$$

Our main result is the following theorem. We use the standard notation $\wedge$ for minimum and $\vee$ for maximum.

Theorem 1 *The classical capacity of a quantum channel with arbitrary Markovian correlated noise, defined by (3), is given by*

$$C(\Phi) = \lim_{n\to\infty} \sup_{\{p_j^{(n)}, \rho_j^{(n)}\}} \left[\bigwedge_{C \in \mathcal{C}} \bar{\chi}_C^{(n)}(\{p_j^{(n)}, \rho_j^{(n)}\}) \right]. \quad (14)$$

The existence of the limit in (14) is proved in Lemma 22 of Appendix A.

Before proving Theorem 1, we consider the special case in which the Markov chain has a single communicating class, and the latter is aperiodic and irreducible.

4 Ergodic memory case

In this section we assume that the underlying Markov chain is aperiodic and irreducible (see e.g. [19]) so that in particular, the invariant distribution, $\{\gamma_i\}_{i\in I}$, is unique. It is well-known that the corresponding Markov chain is ergodic and consequently the output states of the channel are also ergodic. In this case, the Markov chain satisfies the property of *convergence to equilibrium*, i.e.,

$$p_{ij}^{(n)} \to \gamma_j \quad \text{as } n \to \infty,$$

where $p_{ij}^{(n)}$ denotes the n-step transition probability from the state i to the state j, $(i, j \in I)$. This implies that the correlation in the noise, acting on successive inputs to the channel, dies out after a sufficiently large number of uses of the channel. Hence, in this case the channel belongs to the class of channels introduced and studied by Kretschmann and Werner [15], and referred to as *forgetful channels*.

Suppose that $\{p_j^{(n)}, \rho_j^{(n)}\}_{j=1}^{J(n)}$ is a sequence of states given by density matrices $\rho_j^{(n)}$ on $\mathcal{H}^{\otimes n}$ with probabilities $p_j^{(n)}$, $\sum_{j=1}^{J(n)} p_j^{(n)} = 1$.

The Holevo quantity for the channel restricted to $\mathcal{A}_n$ is given by

$$\chi(\{p_j^{(n)}, \Phi^{(n)}(\rho_j^{(n)})\}) = S\left(\sum_{j=1}^{J(n)} p_j^{(n)} \Phi^{(n)}(\rho_j^{(n)})\right) - \sum_{j=1}^{J(n)} p_j^{(n)} S(\Phi^{(n)}(\rho_j^{(n)})) \quad (1)$$

The classical capacity of a quantum channel with classical ergodic memory is stated in the following theorem, which is a special case of Theorem 1.

Theorem 2 *The classical capacity of a quantum channel with memory defined by (3), where the underlying Markov chain is aperiodic and irreducible, is given by*

$$\chi^*(\Phi) = \lim_{n\to\infty} \frac{1}{n} \sup_{\{p_j^{(n)}, \rho_j^{(n)}\}} \chi(\{p_j^{(n)}, \Phi^{(n)}(\rho_j^{(n)})\}) \quad (2)$$

The existence of the limit in (2) is proved in Lemma 22 of Appendix A.

This expression for the capacity was in fact stated and proved in [15]. We present an alternative proof which can then be extended to the case of a general Markov chain. The latter is done in Section 5.

The direct part of Theorem 2, i.e., the achievability of any rate $R < \chi^*(\Phi)$, follows from Lemma 1 given below, which is itself a generalization of the Quantum Feinstein Lemma for a memoryless channel [6, 7]. The weak converse part of Theorem 2 is proved in the general case in Section 6.

4.1 Quantum version of Feinstein's Lemma

Lemma 1 *Let Φ_∞ denote a quantum memory channel with Markovian correlated noise, defined by (3). Suppose that the Markov chain is aperiodic and irreducible. Let $\chi^* = \chi^*(\Phi)$ be given by (2). Given $\epsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ there exist at least $N \geq 2^{n(\chi^*-\epsilon)}$ states with density matrices $\tilde{\rho}_1^{(n)}, \dots, \tilde{\rho}_N^{(n)} \in \mathcal{B}(\mathcal{H}^{\otimes n})$, and positive operators $E_1^{(n)}, \dots, E_N^{(n)} \in \mathcal{B}(\mathcal{K}^{\otimes n})$ such that $\sum_{k=1}^N E_k^{(n)} \leq \mathbf{1}^{(n)}$ and*

$$\mathrm{Tr}\left[\Phi^{(n)}\left(\tilde{\rho}_k^{(n)}\right) E_k^{(n)}\right] > 1 - \epsilon. \quad (3)$$

Proof. Choose l_0 so large that

$$\left| \frac{1}{l_0} \sup_{\{p_j^{(l_0)}, \rho_j^{(l_0)}\}} \chi(\{p_j^{(l_0)}, \Phi^{(l_0)}(\rho_j^{(l_0)})\}) - \chi^* \right| \leq \frac{1}{6}\epsilon. \tag{4}$$

Then assume that the supremum is attained for an ensemble $\{p_j^{(l_0)}, \rho_j^{(l_0)}\}_{j=1}^J$, for a finite J.

Denote for $m \in \mathbb{N}$

$$\bar{\sigma}_{ml_0} = \Phi^{(ml_0)}\left(\bar{\rho}_{l_0}^{\otimes m}\right), \tag{5}$$

where

$$\bar{\rho}_{l_0} = \sum_{j=1}^J p_j^{(l_0)} \rho_j^{(l_0)}. \tag{6}$$

These states form a compatible system of states on $\{\mathcal{B}_{ml_0}\}_{m=1}^\infty$ and hence a state $\bar{\phi}_\infty$ on $\mathcal{B}_\infty$ by

$$\bar{\phi}_\infty(A) = \mathrm{Tr}\,(\bar{\sigma}^{ml_0} A) \tag{7}$$

if $A \in \mathcal{B}_{ml_0}$. This state is clearly l_0-periodic, i.e. invariant under translations over multiples of l_0 . Therefore, the mean entropy

$$S_M(\bar{\phi}_\infty) := \lim_{m\to\infty} \frac{1}{m} S\left(\bar{\sigma}^{ml_0}\right) = \inf_{m\in\mathbb{N}} \frac{1}{m} S\left(\bar{\sigma}^{ml_0}\right) \tag{8}$$

exists.

For l_0 sufficiently large, the mean entropy $S_M(\bar{\phi}_\infty)$ is close to $S\left(\bar{\sigma}_{l_0}\right)$, the von Neumann entropy of the average output of l_0 uses of the channel. This is stated in the following lemma.

Lemma 2 *Given $\epsilon > 0$ there exists $L > 0$ such that for $l_0 \geq L$,*

$$\left| \frac{1}{l_0} S_M(\bar{\phi}_\infty) - \frac{1}{l_0} S(\Phi^{(l_0)}(\bar{\rho}_{l_0})) \right| < \frac{\epsilon}{8} \tag{9}$$

Here $\bar{\phi}_\infty$ is given by (7). The proof is similar to that of Lemma 2.

Henceforth l_0 is fixed to a value such that Lemma 2 and (4) hold. For notational simplicity, explicit dependence on l_0 is often suppressed.

The proof of Lemma 1 requires the sequence of lemmas given below.

Lemma 3 *The state $\bar{\phi}_\infty$ is strongly clustering and hence completely ergodic for l_0-shifts, i.e., for any $A, B \in \mathcal{B}_{ml_0}$,*

$$\lim_{k\to\infty} \bar{\phi}_\infty(A\,\tau^{kl_0}(B)) = \mathrm{Tr}(\bar{\sigma}_{ml_0}\,A)\ \mathrm{Tr}(\bar{\sigma}_{ml_0}\,B). \tag{10}$$

Proof The proof is standard and relies on the fact that the expectations of A and B in the state $\bar{\phi}_\infty$ decouple as their supports are separated by a sufficiently large distance. This is because

$$\sum_{i_2,i_3,\ldots,i_k} q_{i_1 i_2}\ldots q_{i_{k-1}i_k} g(i_k) \to \sum_i \gamma_i g(i), \tag{11}$$

as $k \to \infty$, for any function $g(i)$, since the Markov chain is irreducible and aperiodic. $\square$

In the following we denote $\mathcal{K}^{\otimes l_0}$ by $\mathcal{K}_{l_0}$. We also use the following lemma, which is proved in Appendix B.

Lemma 4 *For any $\delta > 0$, there exists $m_1 \in \mathbb{N}$ such that for all $m \geq m_1$ there exists a subspace $\mathcal{T}_\epsilon^{(m)} \subset \mathcal{K}_{l_0}^{\otimes m}$ with projection $\bar{P}_{ml_0}$ such that*

$$\bar{P}_{ml_0}\,\bar{\sigma}_{ml_0}\,\bar{P}_{ml_0} \leq 2^{-m[S_M(\bar{\phi}_\infty)-\frac{1}{4}\epsilon]}\mathbf{1}^{(ml_0)} \tag{12}$$

and

$$\mathrm{Tr}\left(\bar{\sigma}_{ml_0}\,\bar{P}_{ml_0}\right) > 1-\delta^2. \tag{13}$$

Here $\mathbf{1}^{(ml_0)}$ denotes the identity operator in $\mathcal{B}(\mathcal{K}^{\otimes ml_0})$.

In order to obtain the first term in the expression (2) for the capacity, we need to be able to replace $S_M(\bar{\phi}_\infty)$ in the above lemma by $S(\bar{\sigma}_{l_0})$. This is possible due to Lemma 2.

We need an analogous result to Lemma 4 for the second term in the expression (2) of $\chi^*(\Phi)$. This is stated in Lemma 6 (which is proved in Appendix C). It uses Lemma 5, given below. To formulate these lemmas, we define density matrices Σ_{ml_0} in algebras

$$\mathcal{M}_{ml_0} = \bigoplus_{j_1,\ldots,j_m=1}^{J} \mathcal{B}(\mathcal{K}_{l_0}^{\otimes m})$$

by

$$\Sigma_{ml_0} = \bigoplus_{j_1,\dots,j_m} p_{\underline{j}}^{(m)} \Phi^{(ml_0)} \left(\rho_{j_1}^{(l_0)} \otimes \cdots \otimes \rho_{j_m}^{(l_0)} \right), \tag{14}$$

where $p_{\underline{j}}^{(m)} = \prod_{\alpha=1}^{m} p_{j_\alpha}^{(l_0)}$ and $\rho_j^{(l_0)}$, $j \in \{1, 2, \dots, J\}$, belongs to the maximising ensemble (c.f. (4)). In the following we denote $\rho_{\underline{j}}^{(ml_0)} = \rho_{j_1}^{(l_0)} \otimes \cdots \otimes \rho_{j_m}^{(l_0)}$, with $\underline{j} = (j_1, j_2, \dots, j_m)$, for any $m \in \mathbb{N}$.

Lemma 5 *There exists a unique translation-invariant state ψ_∞ on $\mathcal{M}_\infty = \overline{\bigcup_{m=1}^{\infty} \mathcal{M}_{ml_0}}$ such that*

$$\psi_\infty(A) = \mathrm{Tr}\left(\Sigma_{ml_0} A\right) \tag{15}$$

for $A \in \mathcal{M}_{ml_0}$. Moreover, this state is strongly clustering and therefore completely ergodic.

Proof. The proof of this lemma is similar to that of Lemma 3. $\square$

Note that the mean entropy of ψ_∞ is given by

$$S_M(\psi_\infty) = \lim_{k\to\infty} \frac{1}{k} S(\Sigma_{kl_0}), \tag{16}$$

where

$$\begin{aligned} S(\Sigma_{kl_0}) &= \sum_{j_1,\dots,j_k} S\left(p_{\underline{j}}^{(k)} \Phi^{(kl_0)} (\rho_{j_1}^{(l_0)} \otimes \cdots \otimes \rho_{j_k}^{(l_0)}) \right) \\ &= \sum_{\underline{j}} p_{\underline{j}}^{(k)} S\left(\Phi^{(kl_0)} (\rho_{\underline{j}}^{(kl_0)}) \right) + k\, H\left(\{p_j^{(l_0)}\}_{j=1}^{J} \right). \end{aligned} \tag{17}$$

We define

$$\bar{S}_M \equiv \bar{S}_M\left(\{p_j^{(l_0)}, \rho_j^{(l_0)}\} \right) := \lim_{k\to\infty} \frac{1}{k} \sum_{\underline{j}} p_{\underline{j}}^{(k)} S\left(\Phi^{(kl_0)} (\rho_{\underline{j}}^{(kl_0)}) \right). \tag{18}$$

Lemma 6 *Given $\delta > 0$, there exists $m_2 \in \mathbb{N}$ such that for all $m \geq m_2$ there exist, for all $\underline{j} = (j_1, \dots, j_m) \in \{1, \dots, J\}^m$, one-dimensional subspaces $\mathcal{T}_{\underline{j},\underline{k}}^{(m)}$ of $\mathcal{K}_{l_0}^{\otimes m}$ (indexed by $\underline{k}$ in some set $T_{\underline{j},\epsilon}^{(m)}$) with projections $\pi_{\underline{j},\underline{k}}^{(ml_0)}$ in the $\underline{j}$-th component of $\mathcal{M}_{ml_0}$, such that for all $\underline{k} \in T_{\underline{j},\epsilon}^{(m)}$,*

$$\left| \frac{1}{m} \log \omega_{\underline{j},\underline{k}}^{(ml_0)} + \bar{S}_M(\{p_j^{(l_0)}, \rho_j^{(l_0)}\}) \right| < \frac{\epsilon}{4}, \tag{19}$$

where $\omega_{\underline{j},\underline{k}}^{(ml_0)} = \mathrm{Tr}\left(\Phi^{(ml_0)}(\rho_{\underline{j}}^{(ml_0)})\pi_{\underline{j},\underline{k}}^{(ml_0)}\right)$*, and*

$$\psi_\infty\left(\bigoplus_{\underline{j}}\bigoplus_{\underline{k}\in T_{\underline{j},\epsilon}^{(m)}}\pi_{\underline{j},\underline{k}}^{(ml_0)}\right) > 1-\delta^2. \tag{20}$$

Proof See Appendix C.

We now continue the proof of the theorem. In the following we denote

$$P_{\underline{j}}^{(ml_0)} = \bigoplus_{\underline{k}\in T_{\underline{j},\epsilon}^{(m)}}\pi_{\underline{j},\underline{k}}^{(ml_0)} \tag{21}$$

The remainder of the proof is in fact analogous to that for the case of a memoryless channel (see [6], [7]), so we only sketch the main steps.

For arbitrary n, let $m = [n/l_0]$ and denote, $\bar{\Pi}_n = \bar{P}_{ml_0}\otimes\mathbf{1}_{n-ml_0}$, $\Pi_{\underline{j}}^{(n)} = P_{\underline{j}}^{(ml_0)}\otimes\mathbf{1}_{n-ml_0}$ and $\bar{\sigma}^{(n)} = \mathrm{Tr}_{\,(m+1)l_0-n}\bar{\sigma}_{(m+1)l_0}$. Now let $N = N(n)$ be the maximal number of states $\tilde{\rho}_1^{(n)},\dots,\tilde{\rho}_N^{(n)}$ on $\mathcal{H}^{\otimes n}$ for which there exist positive operators $E_1^{(n)},\dots,E_N^{(n)}$ on $\mathcal{K}^{\otimes n}$, of the form $E_j^{(n)} = \mathcal{E}_j^{ml_0}\otimes\mathbf{1}_{n-ml_0}$, such that

(i) $\sum_{k=1}^N E_k^{(n)} \le \bar{\Pi}_n$ and

(ii) $\mathrm{Tr}\,[\,\tilde{\sigma}_k^{(n)}E_k^{(n)}] > 1-\epsilon$ and

(iii) $\mathrm{Tr}\,[\,\bar{\sigma}^{(n)}E_k^{(n)}] \le 2^{-n[\chi^*(\Phi)-\frac{2}{3}\epsilon]}$.

Here $\tilde{\sigma}_k^{(n)} = \Phi^{(n)}(\tilde{\rho}_k^{(n)})$.

For any given $\underline{j}\in\{1,\dots,J\}^m$ define

$$V_{\underline{j}}^{(n)} = \left(\bar{\Pi}_n - \sum_{k=1}^N E_k^{(n)}\right)^{1/2}\bar{\Pi}_n\Pi_{\underline{j}}^{(n)}\bar{\Pi}_n\left(\bar{\Pi}_n - \sum_{k=1}^N E_k^{(n)}\right)^{1/2}. \tag{22}$$

Clearly, $V_{\underline{j}}^{(n)} \le \bar{\Pi}_n - \sum_{k=1}^N E_k^{(n)}$, and we also have:

Lemma 7 *There exists an* $n_1\in\mathbb{N}$ *such that if* $n\ge n_1$ *then*

$$\mathrm{Tr}\,(\bar{\sigma}^{(n)}V_{\underline{j}}^{(n)}) \le 2^{-n[\chi^*(\Phi)-\frac{2}{3}\epsilon]}, \tag{23}$$

for all $\underline{j}$.

Proof. Let $Q_n = \sum_{k=1}^{N(n)} E_k^{(n)}$. Note that Q_n is of the form $Q_n = \tilde{Q}_{ml_0} \otimes \mathbf{1}_{n-ml_0}$, since $E_j^{(n)} = \mathcal{E}_j^{ml_0} \otimes \mathbf{1}_{n-ml_0}$. Note that Q_n commutes with $\bar{\Pi}_n$ by condition (i). Now, by Lemma 4, we have

$$\bar{P}_{ml_0}\bar{\sigma}_{ml_0}\bar{P}_{ml_0} \leq 2^{-m[S_M(\bar{\phi}_\infty)-\frac{1}{4}\epsilon]}\mathbf{1}^{(ml_0)} \tag{24}$$

and, assuming that $l_0 \geq L$, we have by Lemma 2,

$$\begin{aligned}\bar{P}_{ml_0}\bar{\sigma}_{ml_0}\bar{P}_{ml_0} &\leq 2^{-m[S(\Phi^{(l_0)}(\bar{\rho}))-\frac{1}{4}(1+\frac{1}{2}l_0)\epsilon]}\mathbf{1}^{(ml_0)} \\ &\leq 2^{-n[\frac{1}{l_0}S(\Phi^{(l_0)}(\bar{\rho}))-\frac{1}{4}\epsilon]}\mathbf{1}^{(ml_0)}\end{aligned} \tag{25}$$

provided

$$\frac{n-ml_0}{l_0}S(\Phi^{(l_0)}(\bar{\rho})) \leq \frac{1}{4}(n-m-\frac{1}{2}ml_0)\epsilon, \tag{26}$$

which holds if $l_0 \geq 6$ and

$$m \geq \frac{12}{\epsilon}\log\dim\mathcal{K}$$

since $\frac{1}{l_0}S(\Phi^{(l_0)}(\bar{\rho})) \leq \log\dim\mathcal{K}$.

Using this, we get

$$\begin{aligned}&\mathrm{Tr}\,(\bar{\sigma}^{(n)}V_{\underline{j}}^{(n)}) \\ &= \mathrm{Tr}\left[\bar{\sigma}^{(n)}(\bar{\Pi}_n - Q_n)^{1/2}\bar{\Pi}_n\Pi_{\underline{j}}^{(n)}\bar{\Pi}_n(\bar{\Pi}_n - Q_n)^{1/2}\right] \\ &= \mathrm{Tr}\left[\bar{\Pi}_n\bar{\sigma}^{(n)}\bar{\Pi}_n(\bar{\Pi}_n - Q_n)^{1/2}\Pi_{\underline{j}}^{(n)}(\bar{\Pi}_n - Q_n)^{1/2}\right] \\ &= \mathrm{Tr}\left[\bar{P}_{ml_0}\bar{\sigma}_{ml_0}\bar{P}_{ml_0}(\bar{P}_{ml_0}-\tilde{Q}_{ml_0})^{1/2}P_{\underline{j}}^{(ml_0)}(\bar{P}_{ml_0}-\tilde{Q}_{ml_0})^{1/2}\right] \\ &\leq 2^{-n[\frac{1}{l_0}S(\Phi^{(l_0)}(\bar{\rho}))-\frac{1}{4}\epsilon]}\mathrm{Tr}\left[(\bar{P}_{ml_0}-\tilde{Q}_{ml_0})^{1/2}P_{\underline{j}}^{(ml_0)}(\bar{P}_{ml_0}-\tilde{Q}_{ml_0})^{1/2}\right] \\ &\leq 2^{-n[\frac{1}{l_0}S(\Phi^{(l_0)}(\bar{\rho}))-\frac{1}{4}\epsilon]}\mathrm{Tr}\,(P_{\underline{j}}^{(ml_0)}).\end{aligned} \tag{27}$$

However, by Lemma 6,

$$\begin{aligned}\mathrm{Tr}\,(P_{\underline{j}}^{(ml_0)}) &\leq 2^{m[\bar{S}_M(\{p_j^{(l_0)},\rho_j^{(l_0)}\})+\frac{1}{4}\epsilon]} \\ &\leq 2^{n[\frac{1}{l_0}\bar{S}_M(\{p_j^{(l_0)},\rho_j^{(l_0)}\})+\frac{1}{4}\epsilon]} \\ &\leq 2^{n[\frac{1}{l_0}\sum_j p_j^{(l_0)}S(\rho_j^{(l_0)})+\frac{1}{4}\epsilon]},\end{aligned} \tag{28}$$

where the last inequality follows from the subadditivity of the von Neumann entropy. The lemma now follows from (4). $\square$

Since $N(n)$ is maximal it follows that

$$\mathrm{Tr}\,\left(\sigma_{\underline{j}}^{(n)} V_{\underline{j}}^{(n)}\right) \leq 1-\epsilon. \tag{29}$$

Corollary 1

$$\mathbb{E}\left(\mathrm{Tr}\,\left[\sigma_{\underline{j}}^{(n)} V_{\underline{j}}^{(n)}\right]\right) < 1-\epsilon. \tag{30}$$

Lemma 8 *Assume $\eta > 3\delta$. Then for all $n \geq n_2 = m_1 l_0 \vee m_2 l_0$,*

$$\mathbb{E}\left(\mathrm{Tr}\,\left[\sigma_{\underline{j}}^{(n)} \bar{\Pi}_n \Pi_{\underline{j}}^{(n)} \bar{\Pi}_n\right]\right) > 1-\eta. \tag{31}$$

Proof. We write

$$\begin{aligned}
&\mathbb{E}\left(\mathrm{Tr}\,\left[\sigma_{\underline{j}}^{(n)} \bar{\Pi}_n \Pi_{\underline{j}}^{(n)} \bar{\Pi}_n\right]\right) = \\
&= \mathbb{E}\left(\mathrm{Tr}\,\left[\sigma_{\underline{j}}^{(n)} \Pi_{\underline{j}}^{(n)}\right]\right) - \mathbb{E}\left(\mathrm{Tr}\,\left[\sigma_{\underline{j}}^{(n)} (\mathbf{1}-\bar{\Pi}_n) \Pi_{\underline{j}}^{(n)}\right]\right) \\
&\quad -\mathbb{E}\left(\mathrm{Tr}\,\left[\sigma_{\underline{j}}^{(n)} \bar{\Pi}_n \Pi_{\underline{j}}^{(n)} (\mathbf{1}-\bar{\Pi}_n)\right]\right).
\end{aligned} \tag{32}$$

The first term equals $\mathbb{E}\big[\mathrm{Tr}\,\big(\sigma_{\underline{j}}^{(ml_0)} P_{\underline{j}}^{(ml_0)}\big)\big]$, which by Lemma 6 is $> 1-\delta^2$, provided $n \geq n_2$.

Note that

$$\mathbb{E}\left(\mathrm{Tr}\,\left[\sigma_{\underline{j}}^{(n)} (\mathbf{1}-\bar{\Pi}_n) \Pi_{\underline{j}}^{(n)}\right]\right) = \mathbb{E}\left(\mathrm{Tr}\,\left[\sigma_{\underline{j}}^{ml_0} (\mathbf{1}-\bar{P}_{ml_0}) P_{\underline{j}}^{(ml_0)}\right]\right) \tag{33}$$

and similarly

$$\mathbb{E}\left(\mathrm{Tr}\,\left[\sigma_{\underline{j}}^{(n)} \bar{\Pi}_n \Pi_{\underline{j}}^{(n)} (\mathbf{1}-\bar{\Pi}_n)\right]\right) = \mathbb{E}\left(\mathrm{Tr}\,\left[\sigma_{\underline{j}}^{(ml_0)} \bar{P}_{ml_0} P_{\underline{j}}^{(ml_0)} (\mathbf{1}-\bar{P}_{ml_0})\right]\right) \tag{34}$$

Using (33) and (34), the last two terms on the right hand side of (32) can be bounded using Cauchy-Schwarz and Lemma 4 as follows :

$$\mathbb{E}\left(\mathrm{Tr}\,\left[\sigma_{\underline{j}}^{(n)} (\mathbf{1}-\bar{\Pi}_n) \Pi_{\underline{j}}^{(n)}\right]\right) \leq \delta \tag{35}$$

and

$$\mathbb{E}\left(\mathrm{Tr}\,\left[\sigma^{(n)}_{\underline{j}}\bar{\Pi}_n\Pi^{(n)}_{\underline{j}}(\mathbf{1}-\bar{\Pi}_n)\right]\right) \le \delta \tag{36}$$

provided $n \ge n_1$. Choosing $n_3 = n_1 \vee n_2$ and $\delta^2 + 2\delta < \eta$ the result follows.
□

Lemma 9 *Assume $\eta < \frac{1}{3}\epsilon$ and $\eta > 3\delta$. Then for $n \ge n_3 = n_1 \vee n_2$,*

$$\mathrm{Tr}\,\left[\bar{\sigma}_n \sum_{k=1}^{N} E_k\right] = \mathbb{E}\left(\mathrm{Tr}\,\left[\sigma^{(n)}_{\underline{j}} \sum_{k=1}^{N} E_k\right]\right) \ge \eta^2. \tag{37}$$

Proof. Define

$$Q'_n = \bar{\Pi}_n - (\bar{\Pi}_n - Q_n)^{1/2}. \tag{38}$$

By the above corollary,

$$\begin{aligned}
1-\epsilon \;\; &\ge \;\; \mathbb{E}\left\{\mathrm{Tr}\,\left(\sigma^{(n)}_{\underline{j}}(\bar{\Pi}_n - Q'_n)\Pi^{(n)}_{\underline{j}}(\bar{P}_n - Q'_n)\right)\right\} \\
&= \;\; \mathbb{E}\left\{\mathrm{Tr}\,\left(\sigma^{(n)}_{\underline{j}}\bar{\Pi}_n\Pi^{(n)}_{\underline{j}}\bar{\Pi}_n\right)\right\} \\
&\quad -\mathbb{E}\left\{\mathrm{Tr}\,\left(\sigma^{(n)}_{\underline{j}}Q'_n\Pi^{(n)}_{\underline{j}}\bar{\Pi}_n\right) + \mathrm{Tr}\,\left(\sigma^{(n)}_{\underline{j}}\bar{\Pi}_n\Pi^{(n)}_{\underline{j}}Q'_n\right)\right\} \\
&\quad +\mathbb{E}\left\{\mathrm{Tr}\,\left(\sigma^{(n)}_{\underline{j}}Q'_n\Pi^{(n)}_{\underline{j}}Q'_n\right)\right\}.
\end{aligned} \tag{39}$$

Since the last term is positive, we have, by Lemma 8,

$$\mathbb{E}\left\{\mathrm{Tr}\,\left(\sigma^{(n)}_{\underline{j}}Q'_n\Pi^{(n)}_{\underline{j}}P_n\right) + \mathrm{Tr}\,\left(\sigma^{(n)}_{\underline{j}}P_n\Pi^{(n)}_{\underline{j}}Q'_n\right)\right\} \ge \epsilon - \eta > 2\eta. \tag{40}$$

On the other hand, using Cauchy-Schwarz for each term, the left-hand side is bounded by

$$2\left\{\mathbb{E}\left[\mathrm{Tr}\,\left(\sigma^{(n)}_{\underline{j}}Q'^2_n\right)\right]\right\}^{1/2}. \tag{41}$$

Thus,

$$\mathbb{E}\left[\mathrm{Tr}\,\left(\sigma^{(n)}_{\underline{j}}Q'^2_n\right)\right] \ge \eta^2. \tag{42}$$

To complete the proof, we now claim that

$$Q_n \ge (Q'_n)^2. \tag{43}$$

Indeed, this follows on the domain of P_n from the inequality $1-(1-x)^2 \geq x^2$ for $0 \leq x \leq 1$. $\square$

To complete the proof of the theorem, we now have by assumption,

$$\mathrm{Tr}\left[\bar{\sigma}^{(n)} E_k^{(n)}\right] \leq 2^{-n[\chi^*(\Phi)-\frac{2}{3}\epsilon]} \tag{44}$$

for all $k = 1, \ldots, N(n)$. On the other hand, choosing $\eta < \frac{1}{3}\epsilon$ and $\delta < \frac{1}{3}\eta$, we have by Lemma 9,

$$\mathrm{Tr}\left[\bar{\sigma}^{(n)} \sum_{k=1}^{N} E_k\right] \geq \eta^2 \tag{45}$$

provided $n \geq n_3$. It follows that

$$N(n) \geq \eta^2 2^{n[\chi^*(\Phi)-\frac{2}{3}\epsilon]} \geq 2^{n[\chi^*(\Phi)-\epsilon]} \tag{46}$$

for $n \geq n_3$ and $n \geq -\frac{6}{\epsilon}\log\eta$. $\square$

5 The case of a general Markov chain

In the following we write

$$\Phi_{C,i}^{(n)} = \Phi_{i_k} \otimes \cdots \otimes \Phi_{i_{k+n-1}} \tag{47}$$

if $i = i_k$ is in a periodic class C, and where the labelling is modulo the length of the class.

5.1 The direct part of Theorem 1

In this section we prove the direct part of Theorem 1. As in the ergodic case (considered in Section 4), we once again employ a quantum Feinstein Lemma, which is a generalization of Lemma 1 and is given by the following lemma.

Lemma 10 *For all $\epsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that there exist at least $N = N_n \geq 2^{n[C(\Phi)-\epsilon]}$ product states $\tilde{\rho}_1^{(n)}, \ldots, \tilde{\rho}_N^{(n)}$ on $\mathcal{H}^{\otimes n}$ and positive operators $E_1^{(n)}, \ldots, E_N^{(n)}$ on $\mathcal{K}^{\otimes n}$ such that $\sum_{k=1}^{N} E_k^{(n)} \leq \mathbf{1}$ and*

$$\mathrm{Tr}\left[\Phi^{(n)}(\tilde{\rho}_k^{(n)})\, E_k^{(n)}\right] > 1 - \epsilon \tag{48}$$

for all $k = 1, \ldots, N$.

The proof of this lemma is given in Section 5.3. It uses the idea of adding a preamble to the codewords (as was done in [7]) to distinguish between the different classes of the Markov chain. The construction of the preamble is discussed in detail in the following section.

5.2 Construction of a preamble

To distinguish between the different classes, $\Phi_C^{(n)}$, of the quantum channel Φ, we add a preamble to the input state encoding each message in the set $\mathcal{M}_n$. This is given by an m-fold tensor product of suitable states (as described below). Let us first sketch the idea behind adding such a preamble. Helstrøm [11] showed that two states σ_1 and σ_2, occurring with *a priori* probabilities γ_1 and γ_2 respectively, can be distinguished with an asymptotically vanishing probability of error, if a suitable collective measurement is performed on the m-fold tensor products $\sigma_1^{\otimes m}$ and $\sigma_2^{\otimes m}$, for a large enough $m \in \mathbb{N}$. The optimal measurement is projection-valued. The relevant projection operators, which we denote by Π^+ and Π^-, are the orthogonal projections onto the positive and negative eigenspaces of the difference operator $A_m = \gamma_1\sigma_1^{\otimes m} - \gamma_2\sigma_2^{\otimes m}$. Here we generalize this result to distinguish between the different classes $\Phi_C^{(n)}$. If the preamble is given by a state $\omega^{\otimes m}$, then, by using Helstrøm's result, we can construct a POVM which distinguishes between the output states $\sigma_C^{(n)} := \Phi_C^{(n)}(\omega^{\otimes m})$ corresponding to the different classes $\Phi_C^{(n)}$. The outcome of this POVM measurement would in turn serve to determine which class of the channel is being used for transmission.

We first show that there exists a preamble that can distinguish between the different classes, analogous to the branches in [7]. In fact, we want to do more. In the case of periodic classes, we also want to distinguish between initial states of the class. We therefore subdivide the problem into the following four possibilities:

1. To distinguish between two aperiodic classes;
2. To distinguish between an aperiodic class C and an initial state i' of a periodic class C';
3. To distinguish between two periodic classes C and C'; and
4. To distinguish between the states of a single periodic class.

We refer to the aperiodic classes and the periodic classes with given initial state, as *branches* of the channel.

Consider the first problem: distinguishing between two aperiodic classes. We can obviously assume that the $\Phi_C^{(n)} \neq \Phi_{C'}^{(n)}$ for some n: otherwise the classes are identical and we can combine their probabilities. This means that for any pair of aperiodic classes C, C' there exists $n = n(C, C')$ and a state $\omega_{C,C'}^{(n)}$ such that $\Phi_C^{(n)}(\omega_{C,C'}^{(n)}) \neq \Phi_{C'}^{(n)}(\omega_{C,C'}^{(n)})$. In fact, in most cases we can take $n = 1$, and we shall assume this for simplicity in the following, even though this is not necessary.

Introducing the fidelity of two states as in [18],

$$F(\sigma, \sigma') = \text{Trace}\,\sqrt{\sigma^{1/2}\sigma'\,\sigma^{1/2}}, \tag{49}$$

we then have

$$F(\Phi_C^{(1)}(\omega_{C,C'}), \Phi_{C'}^{(1)}(\omega_{C,C'})) \leq f < 1, \tag{50}$$

for all pairs C, C' with $C < C'$ in some arbitrary ordering of $\mathcal{C}_{\text{aper}}$, the set of aperiodic classes.

The following lemma shows that the classes C and C' can be distinguished.

Lemma 11 *For any two aperiodic classes C and C',*

$$F(\Phi_C^{(m)}(\omega_{C,C'}^{\otimes m}), \Phi_{C'}^{(m)}(\omega_{C,C'}^{\otimes m})) \to 0 \text{ as } m \to \infty. \tag{51}$$

Proof. Choose $\alpha > 0$ so small that $1+\alpha < f^{-1}$. First let k be so large that

$$(1-\alpha)\gamma_j < \sum_{i_2,\ldots,i_{k-1}} q_{ii_2}\ldots q_{i_{k-1}j} < (1+\alpha)\gamma_j \tag{52}$$

for all $i,j \in I$. Now let $\{\mathcal{E}_r\}_r$ be a POVM such that

$$F(\sigma,\sigma') = \sum_r \sqrt{\mathrm{Tr}(\sigma\mathcal{E}_r)\ \mathrm{Tr}(\sigma'\mathcal{E}_r)}, \tag{53}$$

(see e.g. Eq.(9.74) in [18]) where we denote

$$\sigma = \Phi_C^{(1)}(\omega_{C,C'}) \text{ and } \sigma' = \Phi_{C'}^{(1)}(\omega_{C,C'}). \tag{54}$$

Then we have

$$\begin{aligned}
&F\left(\Phi_C^{(mk+m)}(\omega_{C,C'}^{\otimes(mk+m)}), \Phi_{C'}^{(mk+m)}(\omega_{C,C'}^{\otimes(mk+m)})\right) \leq \\
&\leq \sum_{r_1,\ldots,r_m} \left[\mathrm{Tr}\left(\Phi_C^{(mk+m)}(\omega_{C,C'}^{\otimes(mk+m)})\bigotimes_{i=1}^m(\mathcal{E}_{r_i}\otimes \mathbf{1}_k)\right)\right. \\
&\qquad\left.\times\ \mathrm{Tr}\left(\Phi_{C'}^{(mk+m)}(\omega_{C,C'}^{\otimes(mk+m)})\bigotimes_{i=1}^m(\mathcal{E}_{r_i}\otimes \mathbf{1}_k)\right)\right]^{1/2} \\
&\leq \sum_{r_1,\ldots,r_m} (1+\alpha)^{m-1}\prod_{i=1}^m \sqrt{\mathrm{Tr}(\sigma\mathcal{E}_{r_i})\ \mathrm{Tr}(\sigma'\mathcal{E}_{r_i})} \\
&= (1+\alpha)^{m-1}F(\sigma,\sigma')^m \to 0.
\end{aligned} \tag{55}$$

□

Next consider the second case, i.e., to distinguish an aperiodic class C and an initial state i' of a periodic class C'. There exists a state $\omega = \omega_{C,i'}$ on $\mathcal{H}$ such that

$$f := F(\Phi_C^{(1)}(\omega_{C,i'}), \Phi_{i'}(\omega_{C,i'})) < 1. \tag{56}$$

Lemma 12 *Let C be an aperiodic class and C' a periodic class with length $L = L(C')$, let $i' \in C'$, and choose $\omega = \omega_{C,i'}$ as above. Then*

$$F\left(\Phi_C^{(m)}(\omega^{\otimes m}), \Phi_{C',i'}^{(m)}(\omega^{\otimes m})\right) \to 0 \tag{57}$$

as $m \to \infty$.

Proof. We proceed as in Lemma 11 and choose $\alpha > 0$ so small that $1+\alpha < f^{-1}$ and let k be so large that (52) holds and in addition such that k is a multiple of L. Again, we let $\{\mathcal{E}_r\}_r$ be a POVM such that

$$F(\sigma,\sigma') = \sum_r \sqrt{\operatorname{Tr}(\sigma\mathcal{E}_r)\ \operatorname{Tr}(\sigma'\mathcal{E}_r)}, \tag{58}$$

where now

$$\sigma = \Phi_C^{(1)}(\omega_{C,i'}) \text{ and } \sigma' = \Phi_{i'}(\omega_{C,i'}). \tag{59}$$

Then

$$\begin{aligned}
& F\left(\Phi_C^{(mk+m)}(\omega_{C,i'}^{\otimes(mk+m)}), \Phi_{i'}^{(mk+m)}(\omega_{C,i'}^{\otimes(mk+m)})\right) \\
\leq\ & \sum_{r_1,\ldots,r_m} \left[\operatorname{Tr}\left(\Phi_C^{(mk+m)}(\omega_{C,C'}^{\otimes(mk+m)}) \bigotimes_{i=1}^m (\mathcal{E}_{r_i}\otimes \mathbf{1}_k)\right)\right. \\
& \qquad \left. \times \prod_{j=1}^m \operatorname{Tr}\left(\Phi_{i'}(\omega_{C,i'})\,\mathcal{E}_{r_j}\right)\right]^{1/2} \\
\leq\ & \sum_{r_1,\ldots,r_m} (1+\alpha)^{\frac{m-1}{2}} \prod_{j=1}^m \sqrt{\operatorname{Tr}(\sigma\mathcal{E}_{r_j})\ \operatorname{Tr}(\sigma'\mathcal{E}_{r_j})} \\
=\ & (1+\alpha)^{\frac{m-1}{2}} F(\sigma,\sigma')^m \to 0 \quad \text{as } m\to\infty.
\end{aligned} \tag{60}$$

□

Distinguishing two periodic classes is straightforward:

Lemma 13 *If C and C' are two different periodic classes with periods $L(C)$ and $L(C')$ respectively, then there exists a state $\omega_{C,C'}^{(L)}$ on $\mathcal{H}^{\otimes L}$, where $L = L(C)\,L(C')$ such that*

$$F\left(\Phi_C^{(mL)}((\omega_{C,C'}^{(L)})^{\otimes m}), \Phi_{C'}^{(mL)}((\omega_{C,C'}^{(L)})^{\otimes m})\right) \to 0 \tag{61}$$

as $m \to \infty$.

Proof. Since the two periodic classes are distinct, there exists a state $\omega = \omega_{C,C'}^{(L)}$ such that

$$\sigma = \Phi_C^{(L)}(\omega_{C,C'}^{(L)}) \neq \Phi_{C'}^{(L)}(\omega_{C,C'}^{(L)}) \tag{62}$$

(In fact we can take L to be the least common multiple of $L(C)$ and $L(C')$.) Then writing $\omega = \omega_{C,C'} \otimes \varphi^{\otimes k}$, where φ is an arbitrary state on $\mathcal{H}$ and k is so large that (52) holds,

$$\begin{aligned} & F\left(\Phi_C^{(mL+mk)}(\omega^{\otimes m}), \Phi_{C'}^{(mL)}(\omega^{\otimes m})\right) \leq \\ \leq \quad & (1+\alpha)^m F\left(\left(\Phi_C^{(L)}(\omega)\right)^{\otimes m}, \left(\Phi_{C'}^{(L)}(\omega)\right)^{\otimes m}\right) \to 0. \end{aligned} \tag{63}$$

□

Finally, to distinguish the initial states of a given periodic class C, notice first of all that the corresponding CPT maps Φ_i *need not all be distinct!* However, we may assume that there is no internal periodicity of these maps within a periodic class; otherwise the class can be contracted to a single such period. This means, that for any two states $i, i' \in C$ there exists $l \leq L(C)-1$ such that $\Phi_{i+l} \neq \Phi_{i'+l}$. Then choose $\omega = \omega_{i,i'}$ such that

$$f := F(\Phi_{i+l}(\omega), \Phi_{i'+l}(\omega)) < 1. \tag{64}$$

Lemma 14 *If C is a periodic class with period $L(C)$, $i, i' \in C$ and ω is a state as above, then*

$$F\left(\Phi_{C,i}^{(m)}(\omega^{\otimes m}), \Phi_{C,i'}^{(m)}(\omega^{\otimes m})\right) \to 0 \quad \text{as } m \to \infty \tag{65}$$

Proof.

$$\begin{aligned} & F\left(\Phi_{C,i}^{(m)}(\omega^{\otimes m}), \Phi_{C,i'}^{(m)}(\omega^{\otimes m})\right) \\ = \quad & \left[F\left(\Phi_{C,i}^{(L)}(\omega^{\otimes L}), \Phi_{C,i'}^{(L)}(\omega^{\otimes m})\right)\right]^m \\ \leq \quad & [F(\Phi_{i+l}(\omega), \Phi_{i'+l}(\omega))]^m = f^m \to 0. \end{aligned} \tag{66}$$

□

We now introduce, in each of the four cases, difference operators $A_{C,C'}^{(m)}$, $A_{C,i'}^{(m)}$, $A_{i,i'}^{(m)}$ with i, i' in a periodic class, and corresponding projections $\Pi_{C,C'}^{\pm}$, $\Pi_{C,i'}^{\pm}$ and $\Pi_{i,i'}^{\pm}$ onto their positive and negative eigenspaces , which serve to distinguish the different possibilities, as in [7]. The difference operators are defined by

$$A^{(m)}_{C,C'} = \gamma_C \big(\Phi^{(m)}_C(\omega_{C,C'})\big)^{\otimes m} - \gamma_{C'} \big(\Phi^{(m)}_{C'}(\omega_{C,C'})\big)^{\otimes m}, \tag{67}$$

$$A^{(m)}_{C,i'} = \gamma_C \big(\Phi^{(m)}_C(\omega_{C,i'})\big)^{\otimes m} - \gamma_{i'} \big(\Phi^{(m)}_{C'}(\omega_{C,i'})\big)^{\otimes m}, \tag{68}$$

and

$$A^{(m)}_{i,i'} = \gamma_i \big(\Phi^{(m)}_{C,i'}(\omega_{i,i'})\big)^{\otimes m} - \gamma_{i'} \big(\Phi^{(m)}_{C,i'}(\omega_{i,i'})\big)^{\otimes m}. \tag{69}$$

The following lemma was proved in [7]:

Lemma 15 *Suppose that for a given $\delta > 0$,*

$$|\mathrm{Tr}\,[|A^{(m)}_{c,c'}|] - (\gamma_c + \gamma_{c'})| \leq \delta. \tag{70}$$

Then

$$|\mathrm{Tr}\,[\Pi^{+}_{c,c'}\big(\Phi^{(m)}_c(\omega^{\otimes m}_{c,c'})\big)] - 1| \leq \frac{\delta}{2\gamma_c} \tag{71}$$

and

$$|\mathrm{Tr}\,[\Pi^{-}_{c,c'}\big(\Phi^{(m)}_{c'}(\omega^{\otimes m}_{c,c'})\big)] - 1| \leq \frac{\delta}{2\gamma_{c'}}. \tag{72}$$

Here c, c' denote either two different classes C, C' or one aperiodic class C and an initial state i' in a periodic class, or two different initial states in the same periodic class.

To compare the outputs of all the different branches of the channel, we define projections $\tilde{\Pi}_i$ on the tensor product space $\mathcal{K}^{\otimes mM}$ where

$$M = M_1 + M_2 + M_3 + M_4, \tag{73}$$

with

1. M_1 is the total number of pairs of aperiodic classes;

2. M_2 is the total number of pairs of periodic classes;

3. M_3 is the total number of pairs of aperiodic classes and intial states of periodic classes and

4. M_4 is the total number of pairs of states in the same periodic class.

We introduce an arbitrary order on the classes $C \in \mathcal{C}$ assuming $C < C'$ if $C \in \mathcal{C}_{\mathrm{aper}}$ and $C' \in \mathcal{C}_{\mathrm{per}}$. Then we put

$$\tilde{\Pi}_c = \bigotimes_{(c',c''):\, c'<c''} \Gamma^{(c)}_{c',c''}, \text{ where } \Gamma^{(c)}_{c',c''} = \begin{cases} I_m & \text{if} \quad c' \neq c \text{ and } c'' \neq c \\ \Pi^{-}_{c',c} & \text{if} \quad c'' = c \\ \Pi^{+}_{c,c''} & \text{if} \quad c' = c. \end{cases} \tag{74}$$

It follows from the fact that $\Pi^{+}_{c',c''}\Pi^{-}_{c',c''} = 0$, that the projections $\tilde{\Pi}_c$ are also disjoint:

$$\tilde{\Pi}_{c_1}\tilde{\Pi}_{c_2} = 0 \quad \text{for } \ c_1 \neq c_2. \tag{75}$$

We use the following lemma.

Lemma 16 *For all aperiodic classes C,*

$$\lim_{m\to\infty} \mathrm{Tr}\left[\tilde{\Pi}_C\, \Phi^{(mM)}_C\left(\omega^{(mM)}\right)\right] = 1, \tag{76}$$

and for all periodic classes C and all $i \in C$,

$$\lim_{m\to\infty} \mathrm{Tr}\left[\tilde{\Pi}_{C,i}\, \Phi^{(mM)}_{C,i}\left(\omega^{(mM)}\right)\right] = 1. \tag{77}$$

Proof. Notice that for all (c, c'),

$$\begin{aligned} &F(\gamma_c\Phi^{(mM)}_c(\omega_{c,c'})^{\otimes m}, \gamma_{c'}\Phi^{(mM)}_{c'}(\omega_{c,c'})^{\otimes m}) \\ &\qquad = \sqrt{\gamma_c\gamma_{c'}}F(\Phi^{(mM)}_c(\omega_{c,c'}), \Phi^{(mM)}_{c'}(\omega_{c,c'})) \to 0 \end{aligned} \tag{78}$$

as $m \to \infty$. Using the inequalities [18]

$$\begin{aligned} \mathrm{Tr}\,(A_1) + \mathrm{Tr}\,(A_2) - 2F(A_1, A_2) \ &\leq \ ||A_1 - A_2||_1 \\ &\qquad\qquad \leq \mathrm{Tr}\,(A_1) + \mathrm{Tr}\,(A_2) \end{aligned} \tag{79}$$

for any two positive operators A_1 and A_2, we find that

$$\left| \mathrm{Tr}\ \left(|A^{(m)}_{c,c'}|\right) - (\gamma_i + \gamma_j)\right| \leq \delta_m, \tag{80}$$

where $\delta_m \to 0$ as $m \to \infty$, since

$$\mathrm{Tr}\left(|A_{c,c'}^{(m)}|\right) = ||\gamma_c \Phi_c^{(mM)}(\omega_{c,c'})^{\otimes m} - \gamma_{c'} \Phi_{c'}^{(mM)}(\omega_{c,c'})^{\otimes m}||_1. \tag{81}$$

We now replace m by $m' = m + k$, where $k \in \mathbb{N}$ is large enough so that (52) holds, and define

$$\omega^{(m'M)} := \bigotimes_{(c_1,c_2)} \omega_{c_1,c_2}^{\otimes(m+k)}, \tag{82}$$

Using (52) to separate the different classes, we then have for any $C \in \mathcal{C}_{\mathrm{aper}}$,

$$\begin{aligned} 1 \;\geq\; & \mathrm{Tr}\left[\tilde{\Pi}_C \Phi_C^{(m'M)}\left(\bigotimes_{c_1<c_2} \omega_{c_1,c_2}^{\otimes(m+k)}\right)\right] \\ \geq\; & (1-\alpha)^M \prod_{C' \in \mathcal{C}_{\mathrm{aper}}; C'<C} \mathrm{Tr}\left[\Pi_{C',C}^{-}\big(\Phi_C^{(m)}(\omega_{C',C}^{\otimes m})\big)\right] \\ & \times \prod_{C'' \in \mathcal{C}_{\mathrm{aper}}; C''>C} \mathrm{Tr}\left[\Pi_{C,C''}^{+}\big(\Phi_C^{(m)}(\omega_{C,C''}^{\otimes m})\big)\right] \\ & \times \prod_{C' \in \mathcal{C}_{\mathrm{per}}} \prod_{i' \in C'} \mathrm{Tr}\left[\Pi_{C,i'}^{+}\big(\Phi_C^{(m)}(\omega_{C,i'}^{\otimes m})\big)\right] \\ \geq\; & (1-\alpha)^M \left(1 - \frac{\delta_m}{2\gamma_C}\right)^{|\mathcal{C}_{\mathrm{aper}}|-1+\sum_{C' \in \mathcal{C}_{\mathrm{per}}} |C'|} \to 1, \end{aligned} \tag{83}$$

since $\delta_m \to 0$ as $m \to \infty$. The last inequality follows from Lemma 15.

The analogous result, (77), for periodic classes, is proved in a similar manner. □

5.3 Proof of Lemma 10

Given $\delta > 0$, we now fix m_0 so large that

$$\mathrm{Tr}\left[\tilde{\Pi}_C\, \Phi_C^{(m_0 M)}\left(\omega^{(m_0 M)}\right)\right] > 1 - \delta \tag{84}$$

for all $C \in \mathcal{C}_{\mathrm{aper}}$ and

$$\mathrm{Tr}\left[\tilde{\Pi}_{C',i'}\, \Phi_{C',i'}^{(m_0 M)}\left(\omega^{(m_0 M)}\right)\right] > 1 - \delta \tag{85}$$

for all $C' \in \mathcal{C}_{\rm per}$ and $i' \in C'$. Here M is given by (73). The product state $\omega^{(m_0 M)}$, defined through (82), is used as a preamble to the input state encoding each message, and serves to distinguish between the different branches of the channel, i.e., between Φ_C, $C \in \mathcal{C}_{aper}$ and $\Phi_{C',i}$, $C' \in \mathcal{C}_{per}$ and $i \in C'$. If $\rho_k^{(n)} \in \mathcal{B}(\mathcal{H}^{\otimes n})$ is a state encoding the k^{th} classical message in the set $\mathcal{M}_n$, then the k^{th} codeword is given by the product state

$$\omega^{(m_0 M)} \otimes \rho_k^{(n)}.$$

We follow the same steps as in the proof of Theorem 5.1 in [7]. First we fix l_0 large enough, and an ensemble $\{p_j^{(l_0)}, \rho_j^{(l_0)}\}$ such that

$$\left| C(\Phi) - \bigwedge_{C \in \mathcal{C}} \bar{\chi}_C^{(l_0)}(\{p_j^{(l_0)}, \rho_j^{(l_0)}\}) \right| < \frac{\epsilon}{6}. \tag{86}$$

As in the ergodic case (Section 4), let $N = \tilde{N}(n)$ be the maximal number of product states $\tilde{\rho}_1^{(n)}, \dots, \tilde{\rho}_N^{(n)}$ on $\mathcal{H}^{\otimes n}$ for which there exist positive operators $E_1^{(n)}, \dots, E_N^{(n)}$ on $\mathcal{K}^{\otimes m_0 M} \otimes \mathcal{K}^{\otimes n}$ such that

(i) $E_k^{(n)} = \sum_{C \in \mathcal{C}_{\rm aper}} \tilde{\Pi}_C \otimes E_{k,C}^{(n)} + \sum_{C' \in \mathcal{C}_{\rm per}} \sum_{i' \in C'} \tilde{\Pi}_{C',i'} \otimes E_{k,i'}^{(n)}$ and $\sum_{k=1}^N E_{k,C}^{(n)} \leq \bar{P}_{C,n}$; $\sum_{k=1}^N E_{k,i'}^{(n)} \leq \bar{P}_{i',n}$ for $i' \in C' \in \mathcal{C}_{\rm per}$, and

(ii) $$\sum_{C \in \mathcal{C}_{\rm aper}} \gamma_C \mathrm{Tr} \left[(\tilde{\Pi}_C \otimes E_{k,C}^{(n)}) \Phi_C^{(m_0 M + n)} \left(\omega^{(m_0 M)} \otimes \tilde{\rho}_k^{(n)} \right) \right] + \sum_{C' \in \mathcal{C}_{\rm per}} \sum_{i' \in C'} \gamma_{i'} \mathrm{Tr} \left[(\tilde{\Pi}_{C',i'} \otimes E_{k,C}^{(n)}) \Phi_{C',i'}^{(m_0 M + n)} \left(\omega^{(m_0 M)} \otimes \tilde{\rho}_k^{(n)} \right) \right] > 1 - \epsilon,$$ and

(iii) $$\sum_{C \in \mathcal{C}_{\rm aper}} \gamma_C \mathrm{Tr} \left[(\tilde{\Pi}_C \otimes E_{k,C}^{(n)}) \Phi_C^{(m_0 M + n)} \left(\omega^{(m_0 M)} \otimes \bar{\rho}^{(n)} \right) \right] + \sum_{C' \in \mathcal{C}_{\rm per}} \sum_{i' \in C'} \gamma_{i'} \mathrm{Tr} \left[(\tilde{\Pi}_{C',i'} \otimes E_{k,i'}^{(n)}) \Phi_C^{(m_0 M + n)} \left(\omega^{(m_0 M)} \otimes \bar{\rho}^{(n)} \right) \right] \leq 2^{-n[C(\Phi) - \frac{1}{2}\epsilon]}.$$

Note that, as in the ergodic case, we can append $\mathbf{1}^{(n - m l_0)}$ to all POVM elements, to reduce the proof to the case $n = m l_0$. In the following we therefore assume $n = m l_0$ for simplicity.

The typical projection $\bar{P}_{C,n}$ for an aperiodic class is defined as before by Lemma 4. For a periodic class C' we define the typical spaces by interlacing those for the product channels $\Phi_i^{\otimes n}$ $(i \in C')$, as follows:

Lemma 17 *Let C' be a periodic class with period L. Given $\epsilon, \delta > 0$, there exists $m'_1 \in \mathbb{N}$ such that for $m \geq m'_1$ there are subspaces $\overline{\mathcal{N}}_{i,\epsilon}^{(n)} \subset \mathcal{K}_{l_0}^{\otimes m}$ $(i \in C')$, $(n = ml_0)$, with projections $\bar{P}_{i,n}$ such that*

$$\bar{P}_{i,n} \Phi_{C',i}(\rho_{l_0}^{\otimes m}) \bar{P}_{i,n} \leq 2^{-m[S_{C'} - \frac{\epsilon}{4}]}, \tag{87}$$

where

$$S_{C'} = \frac{1}{L} \sum_{i=0}^{L-1} S(\Phi_{C',i}^{(l_0)}(\bar{\rho}_{l_0})),$$

and

$$\mathrm{Tr}\left(\Phi_{C',i}^{(n)}(\bar{\rho}_{l_0}^{\otimes m}) \bar{P}_{i,n}\right) > 1 - \delta^2. \tag{88}$$

Proof. We simply let $\overline{\mathcal{N}}_{i,\epsilon}^{(n)}$ be the subspace spanned by the vectors $|\psi_{i,k_1}\rangle \otimes \cdots \otimes |\psi_{i+l_0(n-1),k_n}\rangle$, where $|\psi_{i,k}\rangle$ is an eigenvector of $\Phi_{C',i}^{(l_0)}$ and $|\psi_{i,k_1}\rangle \otimes |\psi_{i,k_{L+1}}\rangle \otimes \cdots \otimes |\psi_{i,k_{[(n-1)/L]L+1}}\rangle$ belongs to the typical space for $\Phi_{C,i}^{(l_0)}(\bar{\rho}_{l_0})$, $|\psi_{i+1,k_2}\rangle \otimes |\psi_{i+1,k_{L+1}}\rangle \otimes \cdots \otimes |\psi_{i+1,k_{[(n-1)/L]L+2}}\rangle$ to that of $\Phi_{C,i+1}^{(l_0)}(\bar{\rho}_{l_0})$, etc. $\square$

Similarly we have:

Lemma 18 *Let C' be a periodic class with period L. Given $i \in C'$, and a sequence $\underline{j} = (j_1, \ldots, j_m) \in \{1, 2, \ldots, J\}^m$, let $P_{i,\underline{j}}^{(n)} = P_{(C',i),\underline{j}}^{(n)}$ be the projection onto the subspace of $\mathcal{K}^{\otimes n}$ spanned by the eigenvectors of*

$$\Phi_{C',i}^{(n)}(\rho_{\underline{j}}^{(l_0)}) = \bigotimes_{r=1}^{m} \Phi_{C',i+(r-1)l_0}^{(l_0)}(\rho_{j_r}^{(l_0)}),$$

with eigenvalues $\lambda_{\underline{j},\underline{k}} = \prod_{r=1}^m \lambda_{i+(r-1)l_0,j_r,k_r}$ such that

$$\left| \frac{1}{m} \log \lambda_{\underline{j},\underline{k}} + \bar{S}_{C'} \right| < \frac{\epsilon}{4}, \tag{89}$$

where

$$\bar{S}_{C'} = \lim_{m\to\infty} \frac{1}{mL} \sum_{i'\in C'} \sum_{\underline{j}} p_{\underline{j}}^{(n)} S\left((\Phi_{i'} \otimes \cdots \otimes \Phi_{i'+ml_0-1})(\rho_{j_1}^{(l_0)} \otimes \cdots \otimes \rho_{j_m}^{(l_0)}) \right). \tag{90}$$

For any $\delta > 0$ there exists m_2' such that for $m \geq m_2'$,

$$\mathbb{E}\left(\mathrm{Tr}\left[\Phi_{C',i}^{(ml_0)} \left(\bigotimes_{r=1}^{m} \rho_{j_r}^{(l_0)} \right) P_{i,\underline{j}}^{(n)} \right] \right) > 1 - \delta^2. \tag{91}$$

Note that $\bar{S}_{C'}$ can be equivalently expressed as

$$\bar{S}_{C'} = \lim_{m\to\infty} \frac{1}{mL} \sum_{i'\in C'} \overline{S}_{i'},$$

with $\overline{S}_{i'}$ as in (13).

The remainder of the proof is identical to that of Theorem 5.1 in [7]. For each $c = C$ or $c = (C', i')$ with $i' \in C' \in \mathcal{C}_{per}$, and $\underline{j} = (j_1, \ldots, j_m)$, we define, as before

$$V_{c,\underline{j}}^{(n)} = \left(\bar{P}_c^{(n)} - \sum_{k=1}^{N} E_{k,c}^{(n)} \right)^{1/2} \bar{P}_c^{(n)} P_{c,\underline{j}}^{(n)} \bar{P}_c^{(n)} \left(\bar{P}_c^{(n)} - \sum_{k=1}^{N} E_{k,c}^{(n)} \right)^{1/2}. \tag{92}$$

Clearly $V_{c,\underline{j}}^{(n)} \leq \bar{P}_c^{(n)} - \sum_{k=1}^{N} E_{k,c}^{(n)}$.

Put

$$V_{\underline{j}}^{(n)} := \sum_{C\in\mathcal{C}_{\mathrm{aper}}} \tilde{\Pi}_C \otimes V_{C,\underline{j}}^{(n)} + \sum_{C'\in\mathcal{C}_{\mathrm{per}}} \sum_{i'\in C'} \tilde{\Pi}_{C',i'} \otimes V_{(C',i'),\underline{j}}^{(n)}. \tag{93}$$

This is a candidate for an additional measurement operator, $E_{N+1}^{(n)}$, for Bob with corresponding input state $\tilde{\rho}_{N+1}^{(n)} = \rho_{\underline{j}}^{(n)} = \rho_{j_1}^{(l_0)} \otimes \rho_{j_2}^{(l_0)} \ldots \otimes \rho_{j_n}^{(l_0)}$. Clearly, the condition (i), given above, is satisfied and we also have

Lemma 19

$$\begin{aligned}
&\sum_{C\in\mathcal{C}_{\mathrm{aper}}} \gamma_C \mathrm{Tr}\left[(\tilde{\Pi}_C \otimes V_{C,\underline{j}}^{(n)}) \Phi_C^{\otimes m'M+n} \left(\omega^{(m'M)} \bar{\rho}_{l_0}^{\otimes[n/l_0]} \right) \right] \\
&\quad + \sum_{C'\in\mathcal{C}_{\mathrm{per}}} \sum_{i'\in C'} \gamma_{i'} \mathrm{Tr}\left[(\tilde{\Pi}_{C',i'} \otimes V_{(C',i),\underline{j}}^{(n)}) \Phi_C^{\otimes m'M+n} \left(\omega^{(m'M)} \otimes \bar{\rho}_{l_0}^{\otimes[n/l_0]} \right) \right] \\
&\qquad \leq 2^{-n[C(\Phi) - \frac{2}{3}\epsilon]},
\end{aligned} \tag{94}$$

with $\gamma_{i'} = 1/L(C')$, *for* $i' \in C' \in \mathcal{C}_{per}$.

Proof. Writing $\bar{\sigma}_C^{(n)} = \Phi_C^{(n)}(\bar{\rho}^{(n)})$, by the proof of Lemma 7, the following inequality holds for an aperiodic class C, for n large enough:

$$\mathrm{Tr}\,(\bar{\sigma}_C^{(n)} V_{C,\underline{j}}^{(n)}) \leq 2^{-n[\bar{\chi}_C - \frac{1}{2}\epsilon]}, \tag{95}$$

where $\bar{\chi}_C = \bar{\chi}_C^{(l_0)}$ is given by (9), for the maximising ensemble, with $n = l_0$ [c.f. (86)].

Then

$$\begin{aligned}
&\sum_{C\in\mathcal{C}_{\rm aper}} \gamma_C \mathrm{Tr}\,\left[(\tilde{\Pi}_C \otimes V_{C,\underline{j}}^{(n)})\,\Phi_C^{(m_0M+n)}\left(\omega^{(m_0M)}\otimes\bar{\rho}^{(n)}\right)\right] \\
&\leq \sum_{C\in\mathcal{C}_{\rm aper}} \gamma_C \mathrm{Tr}\,[\bar{\sigma}_C^{(n)} V_{C,\underline{j}}^{(n)}] \\
&\leq \sum_{C\in\mathcal{C}_{\rm aper}} \gamma_C\, 2^{-n[\bar{\chi}_C - \frac{\epsilon}{2}]}
\end{aligned} \tag{96}$$

where we used the obvious fact that $\tilde{\Pi}_C \leq \mathbf{1}$ and (94).

Similarly, for $i' \in C' \in \mathcal{C}_{per}$, denoting $Q_{n,i'} = \sum_{k=1}^N E_{k,i'}^{(n)}$, we have, using Lemma 17,

$$\bar{P}_{i',n}\Phi_{C',i'}^{(n)}(\rho_{l_0}^{\otimes m})\bar{P}_{i',n} \leq 2^{-m[S_{C'} - \frac{1}{4}\epsilon]}$$

and hence

$$\begin{aligned}
&\mathrm{Tr}\,(\bar{\sigma}_{C',i'}^{(n)} V_{i',\underline{j}}^{(n)}) \\
&= \mathrm{Tr}\,\left[\bar{\sigma}_{C',i'}^{(n)}(\bar{P}_{C',i'}^{(n)} - Q_{n,i'})^{1/2}\bar{P}_{C,i'}^{(n)} P_{i',\underline{j}}^{(n)} \bar{P}_{C',i'}^{(n)}(\bar{P}_{C',i'}^{(n)} - Q_{n,i'})^{1/2}\right] \\
&\leq 2^{-m[S_{C'} - \frac{1}{4}\epsilon]}\mathrm{Tr}\,\left[(\bar{P}_{C',i'}^{(n)} - Q_{n,i'})^{1/2} P_{i',\underline{j}}^{(n)} (\bar{P}_{C',i'}^{(n)} - Q_{n,i'})^{1/2}\right] \\
&\leq 2^{-m[S_{C'} - \frac{1}{4}\epsilon]}\mathrm{Tr}\,(P_{i',\underline{j}}^{(n)}) \\
&\leq 2^{-n[\frac{1}{l_0}(S_{C'} - \bar{S}_{C'}) - \frac{1}{2}\epsilon]} \\
&\leq 2^{-n[\bar{\chi}_{C'}^{(l_0)} - \frac{1}{2}]},
\end{aligned} \tag{97}$$

where (see (11))

$$\bar{\chi}_{C'}^{(l_0)} = \frac{1}{l_0 L}\sum_{i\in C'}\left(S(\Phi_{C',i}^{(l_0)}(\bar{\rho}_{l_0})) - \bar{S}_i\right).$$

In the second last inequality of (97), we use the fact that Tr $(P_{i',\underline{j}}^{(n)}) \leq 2^{m\bar{S}_{C'}+\epsilon/4}$, which is a standard consequence of Lemma 18. We obtain the last line of (97) by using the subadditivity of the von Neumann entropy, as in (28).

Summing (97) over i' and C', and adding to the bound for $C \in \mathcal{C}_{\mathrm{aper}}$, yields the following bound:

$$\begin{aligned}\text{LHS of (94)} \;\leq\; & \sum_{C\in\mathcal{C}_{aper}} \gamma_C 2^{-n[\bar{\chi}_C^{(l_0)}-\frac{\epsilon}{2}]} \\ & + \sum_{C'\in\mathcal{C}_{per}} \sum_{i\in C'} \gamma_{i'} 2^{-n[\bar{\chi}_{C'}^{(l_0)}-\frac{\epsilon}{2}]}\end{aligned} \tag{98}$$

Now by (86),

$$C(\Phi) \leq \bigwedge_{C\in\mathcal{C}} \bar{\chi}_C^{(l_0)} + \frac{\epsilon}{6},$$

and hence

$$2^{-n[\bar{\chi}_C^{(l_0)}-\frac{\epsilon}{2}]} \leq 2^{-n[C(\Phi)-\frac{2}{3}\epsilon]},$$

for all $C \in \mathcal{C}$, and therefore (94) follows. $\square$

By maximality of N it now follows that the condition (ii) above cannot hold and as before we get, upon taking expectations,

Corollary 2

$$\begin{aligned}&\sum_{C\in\mathcal{C}_{\mathrm{aper}}} \gamma_C\, \mathbb{E}\left(\mathrm{Tr}\left[(\tilde{\Pi}_C \otimes V_{C,\underline{j}}^{(n)})\Phi_C^{\otimes m_0M+n}\left(\omega^{(m_0M)} \otimes \tilde{\rho}_k^{(n)}\right)\right]\right) \\ &+ \sum_{C'\in\mathcal{C}_{\mathrm{per}}} \sum_{i'\in C'} \gamma_{i'}\, \mathbb{E}\left(\mathrm{Tr}\left[(\tilde{\Pi}_{C',i'} \otimes V_{i',\underline{j}}^{(n)})\Phi_{C',i'}^{\otimes m_0M+n}\left(\omega^{(m_0M)} \otimes \tilde{\rho}_k^{(n)}\right)\right]\right) \\ &\leq 1-\epsilon.\end{aligned} \tag{99}$$

We also need the following analogue of Lemma 8:

Lemma 20 *Assume $\eta' > 3\delta$. Then, for n large enough,*

$$\begin{aligned}&\sum_{C\in\mathcal{C}_{\mathrm{aper}}} \gamma_C \mathrm{Tr}\left[(\tilde{\Pi}_C \otimes \bar{P}_{C,n} P_{C,\underline{j}}^{(n)} \bar{P}_{C,n})\Phi_C^{\otimes(m_0M+n)}\left(\omega^{(m_0M)} \otimes \rho_{\underline{j}}^{(n)}\right)\right] \\ &+ \sum_{C'\in\mathcal{C}_{\mathrm{per}}} \sum_{i'\in C'} \gamma_{i'}\, \mathrm{Tr}\left[(\tilde{\Pi}_{C',i'} \otimes \bar{P}_{i',n} P_{i',\underline{j}}^{(n)} \bar{P}_{i',n})\Phi_{C',i'}^{\otimes(m_0M+n)}\left(\omega^{(m_0M)} \otimes \rho_{\underline{j}}^{(n)}\right)\right] \\ &> 1-\eta'\end{aligned} \tag{100}$$

Proof. This is a simple consequence of Lemma 8 and its analogue for periodic classes, together with (84) and (85). $\square$

Lemma 21 *Assume $\eta' < \frac{1}{3}\epsilon$ and write*

$$Q_{n,C} = \sum_{k=1}^{N} E_{k,C}^{(n)} \quad (C \in \mathcal{C}_{\text{aper}}) \text{ and } Q_{n,i'} = \sum_{k=1}^{N} E_{k,i'}^{(n)} \quad (i' \in C' \in \mathcal{C}_{\text{per}}). \tag{101}$$

Then for n large enough,

$$\sum_{C \in \mathcal{C}_{\text{aper}}} \gamma_C \mathrm{Tr}\left[(\tilde{\Pi}_C \otimes Q_{n,C}) \Phi_C^{\otimes m'M+n} \left(\omega^{(m'M)} \otimes \rho_{\underline{j}}^{(n)} \right) \right]$$
$$+ \sum_{C' \in \mathcal{C}_{\text{per}}} \sum_{i' \in C'} \gamma_{i'} \mathrm{Tr}\left[(\tilde{\Pi}_{C',i'} \otimes Q_{n,i'}) \Phi_{C',i'}^{\otimes m'M+n} \left(\omega^{(m'M)} \otimes \rho_{\underline{j}}^{(n)} \right) \right] \geq (\eta')^2. \tag{102}$$

Proof. This is analogous to Lemma 9. $\square$

It now follows, as before, that for n large enough, $\tilde{N}(n) \geq (\eta')^2 \, 2^{n[C(\Phi) - \frac{2}{3}\epsilon]}$. We take the following states as codewords:

$$\rho_k^{(m_0M+n)} = \omega^{(m_0M)} \otimes \tilde{\rho}_k^{(n)}. \tag{103}$$

For n sufficiently large we then have

$$N = N_{n+m_0M} = \tilde{N}(n) \geq (\eta')^2 \, 2^{n[C(\Phi) - \frac{2}{3}\epsilon]} \geq 2^{(m_0M+n)[C(\Phi)-\epsilon]}. \tag{104}$$

To complete the proof, we need to show that the set $\{E_k^{(n)}\}_{k=1}^N$ satisfies (48). However, this follows immediately from condition (ii) (after eq.(86)):

$$\begin{aligned}
&\mathrm{Tr}\left[\Phi^{(m_0M+n)} \left(\rho_k^{(m_0M+n)} \right) E_k^{(n)} \right] = \\
&= \sum_{C \in \mathcal{C}_{\text{aper}}} \gamma_C \mathrm{Tr}\left[\Phi_C^{(m_0M+n)} \left(\omega^{(m_0M)} \otimes \tilde{\rho}_k^{(n)} \right) E_k^{(n)} \right] \\
&\quad + \sum_{C' \in \mathcal{C}_{\text{per}}} \sum_{i \in C'} \gamma_i \mathrm{Tr}\left[\Phi_{C',i}^{(m_0M+n)} \left(\omega^{(m_0M)} \otimes \tilde{\rho}_k^{(n)} \right) E_k^{(n)} \right] \\
&= \sum_{C \in \mathcal{C}_{\text{aper}}} \gamma_C \mathrm{Tr}\left[(\tilde{\Pi}_C \otimes E_{k,C}^{(n)}) \, \Phi_C^{(m_0M+n)} \left(\omega^{(m_0M)} \otimes \tilde{\rho}_k^{(n)} \right) \right] \\
&\quad + \sum_{C' \in \mathcal{C}_{\text{per}}} \sum_{i \in C'} \gamma_i \mathrm{Tr}\left[(\tilde{\Pi}_{C',i} \otimes E_{k,i}^{(n)}) \, \Phi_{C',i}^{(m_0M+n)} \left(\omega^{(m_0M)} \otimes \tilde{\rho}_k^{(n)} \right) \right] \\
&> 1 - \epsilon.
\end{aligned} \tag{105}$$

$\square$

6 Proof of the converse part of Theorem 1

In this section we prove that it is impossible for Alice to transmit classical messages reliably to Bob through the channel Φ defined by (3) and (4) at a rate $R > C(\Phi)$. This is the (weak) converse part of Theorem 1, in the sense that the probability of error does not tend to zero asymptotically as the length of the code increases, for any code with rate $R > C(\Phi)$. To prove the weak converse, suppose that Alice encodes messages labelled by $\alpha \in \mathcal{M}_n$ by states $\rho_\alpha^{(n)}$ in $\mathcal{B}(\mathcal{H}^{\otimes n})$. Let the corresponding outputs for the class C of the channel be denoted by $\sigma_{\alpha,C}^{(n)}$, i.e.

$$\sigma_{\alpha,C}^{(n)} = \Phi_C^{(n)}(\rho_\alpha^{(n)}). \tag{106}$$

Further define

$$\bar{\sigma}_C^{(n)} = \frac{1}{|\mathcal{M}_n|} \sum_{\alpha \in \mathcal{M}_n} \sigma_{\alpha,C}^{(n)}. \tag{107}$$

Let Bob's POVM elements corresponding to the codewords $\rho_\alpha^{(n)}$ be denoted by $E_\alpha^{(n)}$, $\alpha = 1, \ldots, |\mathcal{M}_n|$. We may assume that Alice's messages are produced uniformly at random from the set $\mathcal{M}_n$. Then Bob's average probability of error is given by

$$\bar{p}_e^{(n)} := 1 - \frac{1}{|\mathcal{M}_n|} \sum_{\alpha \in \mathcal{M}_n} \text{Tr} \left[\Phi^{(n)}(\rho_\alpha^{(n)}) E_\alpha^{(n)} \right]. \tag{108}$$

We also define the average error corresponding to the class C of the channel as

$$\bar{p}_{e,C}^{(n)} := 1 - \frac{1}{|\mathcal{M}_n|} \sum_{\alpha \in \mathcal{M}_n} \text{Tr} \left[\Phi_C^{\otimes n}(\rho_\alpha^{(n)}) E_\alpha^{(n)} \right], \tag{109}$$

so that

$$\bar{p}_e^{(n)} = \sum_{C \in \mathcal{C}} \gamma_C \bar{p}_{e,C}^{(n)}. \tag{110}$$

Let $X^{(n)}$ be a random variable with a uniform distribution over the set $\mathcal{M}_n$, characterizing the classical message sent by Alice to Bob. Let $Y_C^{(n)}$

be the random variable corresponding to Bob's inference of Alice's message, when the codeword is transmitted through the class C. It is defined by the conditional probabilities

$$\mathbb{P}\,[Y_C^{(n)} = \beta \,|\, X^{(n)} = \alpha] = \mathrm{Tr}\;[\Phi_C^{(n)}(\rho_\alpha^{(n)}) E_\beta^{(n)}]. \tag{111}$$

By Fano's inequality,

$$h(\bar{p}_{e,C}^{(n)}) + \bar{p}_{C,e}^{(n)} \log(|\mathcal{M}_n| - 1) \geq H(X^{(n)} \,|\, Y_C^{(n)}) = H(X^{(n)}) - H(X^{(n)} \,:\, Y_C^{(n)}). \tag{112}$$

Here $h(\cdot)$ denotes the binary entropy and $H(\cdot)$ denotes the Shannon entropy. By the Holevo bound, for $C \in \mathcal{C}_{aper}$ we have

$$\begin{aligned} & H(X^{(n)} \,: Y_C^{(n)}) \\ \leq \;& S\left(\frac{1}{|\mathcal{M}_n|} \sum_{\alpha \in \mathcal{M}_n} \Phi_C^{(n)}(\rho_\alpha^{(n)})\right) - \frac{1}{|\mathcal{M}_n|} \sum_{\alpha \in \mathcal{M}_n} S\left(\Phi_C^{(n)}(\rho_\alpha^{(n)})\right) \\ = \;& n\bar{\chi}_C\left(\left\{\frac{1}{|\mathcal{M}_n|}, \rho_\alpha^{(n)}\right\}_{\alpha \in \mathcal{M}_n}\right), \end{aligned} \tag{113}$$

where $\bar{\chi}_C\left(\left\{\frac{1}{|\mathcal{M}_n|}, \rho_\alpha^{(n)}\right\}_{\alpha \in \mathcal{M}_n}\right)$ is given by (11).

For $C \in \mathcal{C}_{per}$, with period L,

$$\begin{aligned} & H(X^{(n)} \,: Y_C^{(n)}) \\ \leq \;& S\left(\frac{1}{|\mathcal{M}_n|} \sum_{\alpha \in \mathcal{M}_n} \frac{1}{L} \sum_{i \in C} \Phi_{C,i}^{(n)}(\rho_\alpha^{(n)})\right) - \frac{1}{|\mathcal{M}_n|} \sum_{\alpha \in \mathcal{M}_n} S\left(\frac{1}{L} \sum_{i \in C} \Phi_{C,i}^{(n)}(\rho_\alpha^{(n)})\right) \\ = \;& \frac{1}{|\mathcal{M}_n|} \sum_{\alpha \in \mathcal{M}_n} S\big(\frac{1}{L} \sum_{i \in C} \Phi_{C,i}^{(n)}(\rho_\alpha^{(n)}) || \frac{1}{|\mathcal{M}_n|} \sum_{\beta \in \mathcal{M}_n} \frac{1}{L} \sum_{i \in C} \Phi_{C,i}^{(n)}(\rho_\beta^{(n)})\big) \\ \leq \;& \frac{1}{|\mathcal{M}_n| L} \sum_{\alpha \in \mathcal{M}_n} \sum_{i \in C} S\big(\Phi_{C,i}^{(n)}(\rho_\alpha^{(n)}) || \frac{1}{|\mathcal{M}_n|} \sum_{\beta \in \mathcal{M}_n} \Phi_{C,i}^{(n)}(\rho_\beta^{(n)})\big) \\ = \;& \frac{1}{L} \sum_{i \in C} \chi_{C,i}^{(n)}\left(\{\frac{1}{|\mathcal{M}_n|}, \rho_\alpha^{(n)}\}\right) \\ = \;& n\bar{\chi}_C^{(n)}\left(\{\frac{1}{|\mathcal{M}_n|}, \rho_\alpha^{(n)}\}\right). \end{aligned} \tag{114}$$

In the above, we use the convexity of the relative entropy

$$S(\sigma||\omega) := \mathrm{Tr}\,\sigma(\log\sigma - \log\omega),$$

for density matrices σ and ω.

Therefore, for any class C we have the upper bound

$$H(X^{(n)} : Y_C^{(n)}) \leq n\bar{\chi}_C^{(n)} \left(\{ \frac{1}{|\mathcal{M}_n|}, \rho_\alpha^{(n)} \} \right). \tag{115}$$

Inserting this into Fano's inequality, (112), now yields

$$h(\bar{p}_{C,e}^{(n)}) + \bar{p}_{C,e}^{(n)} \log |\mathcal{M}_n| \geq \log |\mathcal{M}_n| - n\,\bar{\chi}_C \left(\left\{ \frac{1}{|\mathcal{M}_n|}, \rho_\alpha^{(n)} \right\}_\alpha \right). \tag{116}$$

However, since

$$C(\Phi) \geq \bigwedge_{C \in \mathcal{C}} \bar{\chi}_C \left(\left\{ \frac{1}{|\mathcal{M}_n|}, \rho_\alpha^{(n)} \right\}_\alpha \right) \tag{117}$$

and $R = \frac{1}{n} \log |\mathcal{M}_n| > C(\Phi)$, there must be at least one class C such that

$$\bar{p}_{e,C}^{(n)} \geq 1 - \frac{C(\Phi) + 1/n}{R} > 0. \tag{118}$$

We conclude from (110) and (118) that

$$\bar{p}_e^{(n)} \geq \left(1 - \frac{C(\Phi) + 1/n}{R} \right) \bigwedge_{C \in \mathcal{C}} \gamma_C. \tag{119}$$

$\square$

Remark

Note that the strong converse property [9, 23] does not hold for general Markovian channels. For example, for a convex combination of memoryless channels[2]:

$$\Phi^{(n)}(\rho^{(n)}) = \sum_{i=1}^{M} \gamma_i \Phi_i^{\otimes n}(\rho^{(n)}), \tag{120}$$

where $\Phi_i : \mathcal{B}(\mathcal{H}) \to \mathcal{B}()$, Bob's error probability does *not* tend to 1 asymptotically in n for a rate R, such that $C(\Phi) < R < \bar{C}(\Phi)$, where

$$\bar{C}(\Phi) := \bigvee_{i=1}^{M} \chi_i^*,$$

and χ_i^* denotes the Holevo capacity [13, 21] of the memoryless channel Φ_i.

[2]A classical version of such a channel was introduced by Jacobs [14] and studied further by Ahlswede [1], who obtained an expression for its capacity.

Acknowledgements

TCD would like to acknowledge the hospitality of the Statistical Laboratory of Cambridge University during his sabbatical. The work was supported by the European Commission through the Integrated Projects SECOQC and FET/QIPC "SCALA".

Appendix A

Lemma 22 *If $\Phi^{(n)}$ is a quantum channel with memory of the form (3). Then the limit in (14) exists. In particular, the limit in (2) exists.*

Proof. Denote

$$\bar{\chi}_n = \sup_{\{p_j^{(n)},\rho_j^{(n)}\}} \bigwedge_{C\in\mathcal{C}} \bar{\chi}_C^{(n)}(\{p_j^{(n)}, \Phi^{(n)}(\rho_j^{(n)})\}). \tag{121}$$

We shall prove that for any $\delta > 0$ there exist n_0 and m_0 such that for all $n' \geq n_0$ and $n \geq m_0 n'$, $\bar{\chi}_n \geq \bar{\chi}_{n'} - \delta$. This proves the lemma because obviously, $0 \leq \bar{\chi}_n \leq \log\ \dim\mathcal{K}$, and it follows that

$$\liminf_{n\to\infty} \bar{\chi}_n \geq \bar{\chi}_{n'} - \delta$$

and hence $\liminf\limits_{n\to\infty} \bar{\chi}_n \geq \limsup\limits_{n'\to\infty} \bar{\chi}_{n'} - \delta$ where $\delta > 0$ is arbitrary.

To prove the statement, let n' be large, and suppose that $\{p_j^{(n')}, \rho_j^{(n')}\}$ is a maximising ensemble for (121), with n replaced by n'. Given $n \geq n'$, put $m = [n/n']$ and $l = n - mn'$. Define the states $\rho_{\underline{j}}^{(n)} = \bigotimes_{r=1}^m \rho_{j_r}^{(n')} \otimes \rho_{j_{m+1}}^{(l)}$, where $\rho_j^{(l)}$ is the reduced state on $\mathcal{H}^{\otimes l}$. Then $\bar{\rho}^{(n)} = \otimes_{r=1}^m \bar{\rho}^{(n')} \otimes \bar{\rho}^{(l)}$, with $\bar{\rho}^{(n')} := \sum_j p_j^{(n')} \rho_j^{(n')}$. We now write for any class $C \in \mathcal{C}$,

$$\begin{aligned}\Phi_C^{(n)}(\bar{\rho}^{(n)}) \quad = \quad & \sum_{i_1,\dots,i_{m+1}\in C}\ \sum_{i'_1,\dots,i'_{m+1}\in C} \frac{q_{i'_1 i_2}}{\gamma_{i_2}} \cdots \frac{q_{i'_m i_{m+1}}}{\gamma_{i_{m+1}}} \\ & \times\sigma_C^{(n')}(i_1, i'_1) \otimes \cdots \otimes \sigma_C^{(n')}(i_m, i'_m) \otimes \sigma_C^{(l)}(i_{m+1}, i'_{m+1}),\end{aligned} \tag{122}$$

where

$$\sigma_C^{(n')}(i,i') = \sum_{i_2,\dots,i_{n'-1}\in C} \gamma_i q_{ii_2} q_{i_2 i_3}\cdots q_{i_{n'-1}i'} \times(\Phi_i\otimes\Phi_{i_2}\otimes\cdots\otimes\Phi_{i'})(\bar\rho^{(n')}) \tag{123}$$

and similarly for $\sigma_C^{(l)}(i,i')$. Let $\gamma = \bigwedge_{i\in I}\gamma_i$. Using positivity of the density operators and the fact that $q_{ij}\le 1\le \gamma_i/\gamma$, we obtain the simple operator inequality

$$\Phi_C^{(n)}(\bar\rho^{(n)}) \le \frac{1}{\gamma^m}\Phi_C^{(n')}(\bar\rho^{(n')})\otimes\cdots\otimes\Phi_C^{(n')}(\bar\rho^{(n')})\otimes\Phi_C^{(l)}(\bar\rho^{(l)}). \tag{124}$$

Inserting this into the definition of $S(\Phi^{(n)}(\bar\rho^{(n)}))$ and using the operator monotonicity of the logarithm and the fact that (γ_i) is the equilibrium distribution, i.e. $\sum_{i\in I}\gamma_i q_{ij}=\gamma_j$, we obtain

$$S(\Phi_C^{(n)}(\rho^{(n)})) \ge mS(\Phi_C^{(n')}(\bar\rho^{(n')})) + S(\Phi_C^{(l)}(\bar\rho^{(l)})) + m\log\gamma. \tag{125}$$

On the other hand, by subadditivity,

$$S(\Phi_C^{(n)}(\rho_{\underline{j}}^{(n)})) \le \sum_{r=1}^m S(\Phi_C^{(n')}(\rho_{j_r}^{(n')})) + S(\Phi_C^{(l)}(\rho_{j_{m+1}}^{(l)})) \tag{126}$$

so that

$$\bar\chi_C^{(n)}\left(\{p_{\underline{j}}^{(n)},\Phi^{(n)}(\rho_{\underline{j}}^{(n)})\}\right) \ge \frac{mn'}{n}\bar\chi_{n'} + \frac{m}{n}\log\gamma, \tag{127}$$

for all $C\in\mathcal{C}$. □

Appendix B

Proof of Lemma 4: Let l_1 be so large that

$$S_M(\bar\phi_\infty) \le \frac{1}{l_1}S(\bar\sigma^{(l_1)}) < S_M(\bar\phi_\infty) + \frac{\epsilon}{8}. \tag{128}$$

Let $\Omega = \{\lambda_k\}$ denote the spectrum of $\bar\sigma^{(l_1)}$, and let π_k be the projection onto the eigenvector with eigenvalue λ_k. For any $r>0$ and $C\subset\mathcal{X}^r$, put

$$q_C = \sum_{(\lambda_{k_1},\dots,\lambda_{k_r})\in C} \pi_{k_1}\otimes\cdots\otimes\pi_{k_r}, \tag{129}$$

and define the probability measures ν_r on Ω^r and ν_∞ on $\Omega^{\mathbb{N}}$ by

$$\nu_r(C) = \mathrm{Tr}(\Phi^{(rl_0l_1)}(\bar{\rho}_{l_0}^{\otimes(rl_1)})q_C) \text{ and } \nu_\infty(C) = \bar{\phi}_\infty(q_C). \tag{130}$$

By Lemma 3, ν_∞ is ergodic and by McMillan's theorem [16] there exists a typical set

$$\begin{aligned} T_\epsilon^{(r)} \;\; = \;\; & \{(\lambda_{k_1}, \ldots, \lambda_{k_r}) \in \Omega^r \mid \\ & \qquad 2^{-r(h_{KS}(\nu_\infty)+\epsilon/8)} \le \nu_r(\{(\lambda_{k_1}, \ldots, \lambda_{k_r})\}) \le 2^{-r(h_{KS}(\nu_\infty)-\epsilon/8)}\}\,, \end{aligned} \tag{131}$$

satisfying

$$\nu_r(T_\epsilon^{(r)}) > 1 - \delta^2 \tag{132}$$

for r large enough, where $h_{KS}(\nu_\infty)$ denotes the Kolmogorov-Sinai entropy. Now,

$$h_{KS}(\nu_\infty) = \inf_r \frac{1}{r} H(\nu_r) \le H(\nu_1) = S(\bar{\sigma}^{(l_1)}) < l_1 \left(S_M(\bar{\phi}_\infty) + \frac{\epsilon}{8} \right), \tag{133}$$

where $H(\nu)$ denotes the Shannon entropy corresponding to the probability measure ν. On the other hand

$$h_{KS}(\nu_\infty) \ge l_1\, S_M(\bar{\phi}_\infty) \tag{134}$$

because, by positivity of the relative entropy,

$$\begin{aligned} & S(\bar{\sigma}^{(rl_1)}) \\ & \quad = \;\; -\mathrm{Tr}\left[\bar{\sigma}^{(rl_1)} \log \bar{\sigma}^{(rl_1)}\right] \\ & \quad \le \;\; -\mathrm{Tr}\left[\bar{\sigma}^{(rl_1)} \log \left(\bigoplus_{k_1,\ldots,k_r} \left[\mathrm{Tr}\left(\bar{\sigma}^{(rl_1)}\right](\pi_{k_1} \otimes \cdots \otimes \pi_{k_r})\right) \pi_{k_1} \otimes \cdots \otimes \pi_{k_r} \right)\right] \\ & \quad = \;\; -\sum_{k_1,\ldots,k_r} \mathrm{Tr}[\bar{\sigma}^{(rl_1)} \pi_{k_1} \otimes \cdots \otimes \pi_{k_r}] \log \mathrm{Tr}[\bar{\sigma}^{(rl_1)} \pi_{k_1} \otimes \cdots \otimes \pi_{k_r}] \\ & \quad = \;\; H(\nu_r). \end{aligned} \tag{135}$$

For arbitrary m, let $r = [m/l_1]$ and define

$$\pi_{\underline{k}}^{(m)} = \pi_{k_1} \otimes \cdots \otimes \pi_{k_r} \otimes \mathbf{1} \in \mathcal{B}(\mathcal{K}_{l_0}^{\otimes m}), \quad \underline{k} = (k_1, \ldots, k_r). \tag{136}$$

Let $\overline{T}_\epsilon^{(m)} = \{\underline{k} : (\lambda_{k_1}, \ldots, \lambda_{k_r}) \in T_\epsilon^{(r)}\}$, and define

$$\mathcal{T}_\epsilon^{(m)} = \oplus_{\underline{k} \in \overline{T}_\epsilon^{(m)}} \pi_{\underline{k}}^{(m)} (\mathcal{K}_{l_0}^{\otimes m})$$

. Clearly,

$$\bar{\phi}_\infty \left(\oplus_{\underline{k} \in \overline{T}_\epsilon^{(m)}} \pi_{\underline{k}}^{(m)} \right) = \mathrm{Tr}[\bar{\sigma}^{(rl_1)} q_{T_\epsilon^{(r)}}] = \nu_r \left(T_\epsilon^{(r)} \right) > 1 - \delta^2. \tag{137}$$

Moreover, if $\underline{k} \in \overline{T}_\epsilon^{(m)}$, it follows from (131), (133) and (134) that

$$\frac{1}{m} \log \nu_r(\{(\lambda_{k_1}, \ldots, \lambda_{k_r})\}) \leq -\frac{rl_1}{m} \left(S_M(\bar{\phi}_\infty) - \frac{1}{l_1} \frac{\epsilon}{8} \right), \tag{138}$$

and

$$\frac{1}{m} \log \nu_r(\{(\lambda_{k_1}, \ldots, \lambda_{k_r})\}) \geq -\frac{rl_1}{m} \left(S_M(\bar{\phi}_\infty) + \left(1 + \frac{1}{l_1} \right) \frac{\epsilon}{8} \right). \tag{139}$$

Taking $l_1 > 3$ and m large enough, we obtain

$$\left| \frac{1}{m} \log \nu_r(\{(\lambda_{k_1}, \ldots, \lambda_{k_r})\}) + S_M(\bar{\phi}_\infty) \right| < \frac{\epsilon}{6}. \tag{140}$$

Now let

$$\bar{P}_{ml_0} = \bigoplus_{\underline{k} \in \overline{T}_\epsilon^{(m)}} \pi_{\underline{k}}^{(m)} \tag{141}$$

and assume that l_1 is so large that $\epsilon l_1 / 12 > -\log \gamma_{\min}$, where $\gamma_{\min} = \bigwedge_{i \in I} \gamma_i$. Note that $\gamma_{\min} > 0$. Define

$$\bar{\sigma}_l(i, i') = \sum_{j_1, \ldots, j_l = 1}^{J} p_{\underline{j}}^{(l)} \sum_{i_2, \ldots, i_{l-1}} \gamma_i q_{ii_2} \ldots q_{i_{l-1} i'} \Phi_i(\rho_{j_1}) \otimes \cdots \otimes \Phi_{i'}(\rho_{j_l}), \tag{142}$$

where $\underline{j} = (j_1, j_2, \ldots, j_l)$. Then we can write as in the proof of Lemma 2,

$$\begin{aligned} \bar{\sigma}_{ml_0} \quad &= \quad \sum_{i_1, \ldots, i_{2r+2}} \frac{q_{i_2 i_3}}{\gamma_{i_3}} \frac{q_{i_4 i_5}}{\gamma_{i_5}} \cdots \frac{q_{i_{2r} i_{2r+1}}}{\gamma_{i_{2r+1}}} \\ &\qquad \times \bar{\sigma}_{l_1}(i_1, i_2) \otimes \cdots \otimes \bar{\sigma}_{l_1}(i_{2r-1}, i_{2r}) \otimes \bar{\sigma}^{(m - rl_1)}(i_{2r+1}, i_{2r+2}). \end{aligned} \tag{143}$$

Using the positivity of the transition probabilities, we have

$$\bar{P}_{ml_0} \bar{\sigma}_{ml_0} \bar{P}_{ml_0} \leq 2^{-m[S_M(\bar{\phi}_\infty) - \frac{\epsilon}{4}]} \mathbf{1}^{(ml_0)}.$$

By the fact that π_k is an eigenprojection of $\bar{\sigma}^{(l_1)}$ we then have

$$\bar{P}_{ml_0}\bar{\sigma}_{ml_0}\bar{P}_{ml_0} \leq \gamma_{\min}^{-r} 2^{-m(S_M(\bar{\phi}_\infty)-\epsilon/6)}\mathbf{1}^{(ml_0)}. \tag{144}$$

But $\gamma_{\min}^{-r} < 2^{-m\epsilon/12}$ by the above assumption. $\square$

Appendix C

Proof of Lemma 6 In the following, we suppress the dependence on l_0. We follow Hiai & Petz [12], as in Lemma 4. Fix $l \geq 12$ large enough so that

$$\frac{1}{l}S(\Sigma_{ll_0}) < S_M(\psi_\infty) - \frac{\epsilon}{12}. \tag{145}$$

Let $\mathcal{Y}_{\underline{j}}^{(l)}$ be the spectrum of $\sigma_{\underline{j}}^{(ll_0)} = \Phi^{(ll_0)}(\rho_{j_1}^{(l_0)} \otimes \rho_{j_2}^{(l_0)} \ldots \otimes \rho_{j_l}^{(l_0)})$. Note that Σ_{ll_0} can be represented as a block-diagonal matrix in $\bigoplus_{j_1,\ldots,j_l=1}^{J} \mathcal{K}_{l_0}^{\otimes l}$ with spectrum consisting of eigenvalues $\nu_{\underline{j},k} = p_{\underline{j}}^{(l)}\alpha_{\underline{j},k}$ with $\underline{j} \in \{1,\ldots,J\}^l$, $k = 1,\ldots,(\dim(\mathcal{K}_{l_0}))^l$, and $\alpha_{\underline{j},k}$ being the eigenvalues of $\sigma_{\underline{j}}^{(ll_0)}$. Let

$$\mathcal{Y}_l = \bigcup_{\underline{j}\in\{1,\ldots,J\}^l} \mathcal{Y}_{\underline{j}}^{(l)}. \tag{146}$$

We now define measures μ_s, for $s \in \mathbb{N}$, on $(\mathcal{Y}_l)^s$ by

$$\mu_s(C) = \sum_{\underline{j}\in\{1,\ldots,J\}^{sl}} p_{\underline{j}}^{(sl)} \operatorname{Tr}\left(\sigma_{\underline{j}}^{(sl)} q_C^{(s)}\right), \tag{147}$$

where $C \subset (\mathcal{Y}_l)^s$, and

$$q_C^{(s)} = \sum_{(\lambda_{\underline{j}_1,k_1},\ldots,\lambda_{\underline{j}_s,k_s})\in C} \pi_{\underline{j}_1,k_1} \otimes \cdots \otimes \pi_{\underline{j}_s,k_s} \tag{148}$$

for $\underline{j} = (\underline{j}_1,\ldots,\underline{j}_s)$. (Here $\pi_{\underline{j},k}$ denotes the projection onto the k-th eigenvector of $\sigma_{\underline{j}}^{(l)}$.) We also define the projective limit μ_∞ on $\mathcal{Y}_l^{\mathbb{N}}$ by

$$\mu_\infty(C) = \mu_s(C) = \psi_\infty(q_C^{(s)}), \tag{149}$$

for a cylinder set $C \in (\mathcal{Y}_l)^s$. It follows from Lemma 5 that μ_∞ is ergodic. Define typical sets

$$\begin{aligned}\tilde{T}^{(s)}_{\underline{j},\epsilon} \;=\; & \Big\{(\lambda_{\underline{j}_1,k_1},\ldots,\lambda_{\underline{j}_s,k_s}) \in \mathcal{Y}_l^s \mid \\ & 2^{-s(h_{KS}(\mu_\infty)+\epsilon/12)} \le \mu_s(\{(\lambda_{\underline{j}_1,k_1},\ldots,\lambda_{\underline{j}_s,k_s})\}) \le 2^{-s(h_{KS}(\mu_\infty)-\epsilon/12)}\Big\},\end{aligned} \tag{150}$$

where $h_{KS}(\mu_\infty)$ is the Kolmogorov-Sinai entropy of μ_∞. By McMillan's theorem [16],

$$\mu_s\left(\bigcup_{\underline{j}} \tilde{T}^{(s)}_{\underline{j},\epsilon}\right) > 1 - \frac{1}{2}\delta^2 \tag{151}$$

for s large enough. Now,

$$\begin{aligned} h_{KS}(\mu_\infty) &= \inf_s \frac{1}{s} H(\mu_s) \\ &\le H(\mu_1) = S(\Sigma_l) \\ &< l\left(S_M(\psi_\infty) + \frac{\epsilon}{12}\right), \end{aligned} \tag{152}$$

by (145), and on the other hand

$$h_{KS}(\mu_\infty) \ge l S_M(\psi_\infty) \tag{153}$$

by positivity of the relative entropy.

For arbitrary m we argue as in Lemma 2, and let $s = [m/l]$. Writing, $m = sl + r$, and $\underline{j} = (j_1,\ldots,j_m) = (\underline{j}_1,\ldots,\underline{j}_s,\underline{j}_0)$, we have

$$\pi^{(ml_0)}_{\underline{j},\underline{k}} = \pi_{\underline{j}_1,k_1} \otimes \cdots \otimes \pi_{\underline{j}_s,k_s} \otimes \pi^{(r)}_{\underline{j}_0}, \tag{154}$$

where $\pi^{(r)}_{\underline{j}_0}$ is the projection in $\bigoplus_{j_1,\ldots,j_r=1}^J \mathcal{K}^{\otimes r}$ onto the $\underline{j}_0$-th summand. Let $\tilde{T}^{[m]}_{\underline{j},\epsilon} = \tilde{T}^{(s)}_{\underline{j},\epsilon}$. Then,

$$\begin{aligned} \psi_\infty\left(\bigoplus_{\underline{j}\in\{1,\ldots,J\}^m} \bigoplus_{\underline{k}\in\tilde{T}^{(m)}_{\underline{j},\epsilon}} \pi^{(ml_0)}_{\underline{j},\underline{k}}\right) &= \mathrm{Tr}\left[\Sigma_{sl}\left(\bigoplus_{\underline{j}\in\{1,\ldots,J\}^{sl}} q_{\tilde{T}^{(s)}_{\underline{j},\epsilon}}\right)\right] \\ &= \mu_s\left(\bigcup_{\underline{j}} \tilde{T}^{(s)}_{\underline{j},\epsilon}\right) > 1 - \frac{1}{2}\delta^2. \end{aligned} \tag{155}$$

Moreover, if $(\lambda_{\underline{j}_1,k_1},\ldots,\lambda_{\underline{j}_s,k_s}) \in \tilde{T}^{[m]}_{\underline{j},\epsilon}$,

$$\frac{1}{m}\log\mu_s\left(\{(\lambda_{\underline{j}_1,k_1},\ldots,\lambda_{\underline{j}_s,k_s})\}\right) \le -\frac{sl}{m}\left(S_M(\psi_\infty)-\frac{1}{l}\frac{\epsilon}{12}\right), \tag{156}$$

and

$$\frac{1}{m}\log\mu_s\left(\{(\lambda_{\underline{j}_1,k_1},\ldots,\lambda_{\underline{j}_s,k_s})\}\right) \ge -\frac{sl}{m}\left(S_M(\psi_\infty)+\left(1+\frac{1}{l}\right)\frac{\epsilon}{12}\epsilon\right). \tag{157}$$

Finally define the typical set of indices $\underline{j}$:

$$T^{[m]}_\epsilon = \left\{\underline{j}\in\{1,\ldots,J\}^m \,|\, 2^{-m(H(\{p_j\})+\epsilon/12)} \le p^{(m)}_{\underline{j}} \le 2^{-m(H(\{p_j\})-\epsilon/12)}\right\}. \tag{158}$$

Then for m large enough,

$$\mathbb{P}^{\otimes m}\left[T^{[m]}_\epsilon\right] > 1-\frac{1}{2}\delta^2, \tag{159}$$

if $\mathbb{P}$ denotes the probability with respect to the ensemble probabilities $\{p_j\}_{j=1}^J$. Defining

$$T^{(m)}_{\underline{j},\epsilon} = \begin{cases} T^{[m]}_{\underline{j},\epsilon} & \text{if } \underline{j}\in T^{[m]}_\epsilon \\ \emptyset & \text{if } \underline{j}\notin T^{[m]}_\epsilon, \end{cases} \tag{160}$$

we have for $(\lambda_{\underline{j}_1,k_1},\ldots,\lambda_{\underline{j}_s,k_s}) \in T^{(m)}_{\underline{j},\epsilon}$,

$$\begin{aligned} \frac{1}{m}\log\lambda^{m)}_{\underline{j},\underline{k}} &= -\frac{1}{m}\log(p^{(l)}_{\underline{j}_1}\ldots p^{(l)}_{\underline{j}_s}) + \frac{1}{m}\log\mu_s(\{(\lambda_{\underline{j}_1,k_1},\ldots,\lambda_{\underline{j}_s,k_s})\}) \\ &\le -\frac{sl}{m}\left(S_M(\psi_\infty)-\frac{1}{l}\frac{\epsilon}{12}\right)+H(\{p_j\})+\frac{1}{12}\epsilon \\ &\le -\bar{S}_M+\frac{1}{4}\epsilon, \end{aligned} \tag{161}$$

and

$$\begin{aligned} \frac{1}{m}\log\lambda^{m)}_{\underline{j},\underline{k}} &= -\frac{1}{m}\log(p^{(l)}_{\underline{j}_1}\ldots p^{(l)}_{\underline{j}_s}) + \frac{1}{m}\log\mu_s(\{(\lambda_{\underline{j}_1,k_1},\ldots,\lambda_{\underline{j}_s,k_s})\}) \\ &\ge -\left(\bar{S}_M+\frac{1}{4}\epsilon\right) \end{aligned} \tag{162}$$

for m large enough. Moreover,

$$\begin{aligned}
\psi_\infty\left(\bigoplus_{\underline{j}\in T_\epsilon^{[m]}}\bigoplus_{\underline{k}\in T_{\underline{j},\epsilon}^{(m)}}\pi_{\underline{j},\underline{k}}^{(ml_0)}\right) &= \mathrm{Tr}\left[\Sigma_m\left(\oplus_{\underline{j}\in T_\epsilon^{[m]}} q_{\tilde{T}_{\underline{j},\epsilon}^{(s)}}\right)\right] \\
&= \sum_{\underline{j}\in T_\epsilon^{[m]}} p_{\underline{j}}^{(m)}\,\mathrm{Tr}\left(\Phi^{(m)}(\rho_{\underline{j}}^{(m)})q_{\tilde{T}_{\underline{j},\epsilon}^{(s)}}\right) \\
&\geq \mu_s\left(\bigcup_{\underline{j}}\tilde{T}_{\underline{j},\epsilon}^{(s)}\right) - \mathbb{P}^{\otimes m}\left[\left(T_\epsilon^{[m]}\right)^c\right] \\
&> 1-\delta^2. \qquad (163)
\end{aligned}$$

$\square$

References

[1] R. Ahlswede, "The Weak Capacity of Averaged Channels", Z. Wahrscheinlichkeitstheorie verw. Geb. **11**, 61–73 (1968).

[2] I. Bjelaković and H. Boche, "Ergodic Classical-Quantum Channels: Structure and Coding Theorems", *quant-ph/0609229.*

[3] I. Bjelaković and H. Boche, "Classical capacities of compound and averaged quantum channels", *arXiv:0710.3027.*

[4] G. Bowen and S. Mancini, "Quantum channels with a finite memory", *Phys. Rev. A* **69**, 01236, 2004.

[5] T. M. Cover and J. A. Thomas, *Elements of Information Theory,* John Wiley & Sons, Inc.

[6] N. Datta and T. C. Dorlas, "A Quantum Version of Feinstein's Lemma and its application to Channel Coding", *Proc. of Int. Symp. Inf. Th. ISIT 2006, Seattle,* 441-445 (2006).

[7] N. Datta and T. C. Dorlas, "The coding theorem for a class of channels with long-term memory", *J. Phys. A: Math. Theor.* **40**, 8147–8164 (2007).

[8] A. Feinstein, “ A new basic theorem of information theory,” *IRE Trans. PGIT,* **4**, pp. 2–22, 1954.

[9] M. Hayashi and H. Nagaoka, “General formulas for capacity of classical-quantum channels,” *IEEE Trans. Inform. Theory* **49**, pp. 1753–1768, 2003.

[10] A. I. Khinchin, *Mathematical Foundations of Information Theory,* Dover Publications, 1957. Part II: On the Fundamental Theorems of Information Theory, Chapter IV.

[11] C. W. Helstrøm, *Quantum Detection and Estimation Theory,* Mathematics in Science and Engineering, **vol. 123**, Academic Press, London 1976.

[12] F. Hiai & D. Petz, “The proper formula for the relative entropy and its asymptotics in quantum probability”. *Commun. Math. Phys.* **143**, 257–281, 1991.

[13] A. S. Holevo, “The capacity of a quantum channel with general signal states,” *IEEE Trans. Info. Theory*, **44**, 269-273, 1998.

[14] K.Jacobs, “Almost periodic channels.” Colloqium on Comb. Methods in Prob. Theory. Aarhus 1962.

[15] D. Kretschmann and R. F. Werner, “Quantum channels with memory,” Phys. Rev. A **72**, 062323, 2005; *quant-ph/0502106.*

[16] B. McMillan. The basic theorems of information theory. *Ann. Math. Stat.* **24**, 196–219, 1953.

[17] C. Macchiavello and G. M. Palma, “Entanglement–enhanced information transmission over a quantum channel with correlated noise”, *Phys. Rev. A* **65**, 050301, 2002.

[18] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information,* Cambridge University Press, Cambridge, 2000.

[19] J. R. Norris, *Markov Chains, Cambridge Series in Statistical and Probabilistic Mathematics* Cambridge University Press, Cambridge, 1997.

[20] M. Ohya and D. Petz, *Quantum Entropy and Its Use*, Springer-Verlag, 1993.

[21] B. Schumacher and M. D. Westmoreland, "Sending classical information via noisy quantum channels," *Phys. Rev. A* **56**, 131-138, 1997.

[22] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, Part I, pp. 379–423, 1948; Part II, pp. 623–656, 1948.

[23] A. Winter, "Coding theorem and strong converse for quantum channels," *IEEE Trans. Info. Theory*, **45**, 2481–2485, 1999.